



UNIVERSIDAD COMPLUTENSE DE MADRID



FACULTAD DE INFORMÁTICA

SISTEMAS INFORMÁTICOS

CURSO 2009/2010

# SISTEMA DE ANÁLISIS DE ENCUESTAS DE MARCAS DE DISTRIBUIDOR

**Jorge Caballero Prados**

**Pablo Rodrigo Jimeno Hidalgo**

**Raúl Ortega Maestro**

*Dirigido por:*

***Dra. María Victoria López López***

*Departamento de Arquitectura de Computadores y Automática*

*Facultad de Informática - Universidad Complutense de Madrid*



## **Autorización**

Nosotros, los autores del proyecto Sistema de Análisis de Encuestas de Marcas de Distribuidor, de la asignatura de Sistemas Informáticos autorizamos a la Universidad Complutense de Madrid a utilizar y difundir con fines académicos, no comerciales, el contenido de este documento de texto, así como del contenido del CD complementario que adjuntamos con él mismo.



Jorge Caballero Prados  
DNI: 50210766-H.



Pablo Rodrigo Jimeno Hidalgo  
DNI: 51453254-T



Raúl Ortega Maestro  
DNI: 53617711-L

Dirigidos por:

Dra. María Victoria López López.  
*Departamento de Arquitectura de Computadores y Automática  
Facultad de Informática - Universidad Complutense de Madrid*





## **Agradecimientos**

Queremos expresar nuestro agradecimiento a todos los colaboradores que han participado, ayudándonos en el desarrollo de este proyecto:

*A los responsables de desarrollo de SPSS, en especial a su ingeniero jefe John Peck, por responder rápida y precisamente, a través de su foro, <http://www.spss.com/fusetalk/forum/>, a todas las dudas que nos han asaltado durante la realización de este proyecto.*

*A Juan Carlos Martínez Sánchez por ayudarnos a elegir el lenguaje y programa para este proyecto.*

*A los profesores de los diferentes cursos a los que hemos asistido, por hacernos más fácil y cómoda la realización de este proyecto.*

*A nuestra tutora, María Victoria López López, por aconsejarnos durante todo el proyecto y, en especial, a ayudarnos en la búsqueda de información y elaboración del artículo.*



## Índice

|  |   |        |
|--|---|--------|
| Prólogo  | í     | ..9    |
| Resumen  | í     | .11    |
| Abstract   | í     | ....11 |
| Introducción   | ..í ..í | ....13 |
| 1. Estado del arte   | .....   | 15     |
| 1.1. Introducción  | .....   | 15     |
| 1.2. Software estadístico  | .....   | 16     |
| 2. Conceptos teóricos previos  | .....   | 23     |
| 2.1. Marcas de Distribuidor  | .....   | 23     |
| 2.2. Etapas evolutivas en las MDD  | .....   | 25     |
| 2.3. Análisis Estadístico de MDD   | .....   | 29     |
| 2.4. Técnicas Soft Computing aplicadas a MDD                                     | .....   | 40     |
| 3. Desarrollo del proyecto   | .....   | 47     |
| 3.1. Introducción  | .....   | 47     |
| 3.2. Tecnologías Utilizadas  | .....   | 51     |
| 3.3. Diseño y modelado de las Bases de Datos                                     | .....   | 59     |
| 3.4. Producto Final  | .....   | 67     |
| 3.4.1. Módulo 1: Visualización de los datos                                      | .....   | 68     |
| 3.4.2. Módulo 2: Análisis de encuestas de consumidores                           | .....   | 70     |
| 3.4.3. Módulo 3: Análisis de Independencia                                       | .....   | 76     |
| 3.4.4. Modulo 4: Comparación de consumo usando métodos de decisión multicriterio | .....   | 81     |
| 3.4.5. Módulo 5: Gráficos  | .....   | 83     |
| 3.4.6. Módulo 6: Ayuda   | .....   | 85     |
| 4. Manual de uso   | .....   | 89     |
| 4.1. Herramientas de autor   | .....   | 89     |
| 4.2. Manual de uso y documento e-Learning  | .....   | 94     |
| 5. Conclusiones y trabajo futuro   | .....   | 95     |
| 6. Bibliografía  | .....   | 97     |
| 7. Anexo   | .....   | 99     |



## Prólogo

Esta memoria está realizada por alumnos de la asignatura Sistemas Informáticos, por lo tanto, constituye su Proyecto Fin de Carrera. En ella, los alumnos muestran el trabajo realizado a partir de una gran base de datos procedente del Departamento de Comercialización e Investigación de Mercados de esta Universidad Complutense. El trabajo consiste en organizar los datos mediante diseño y administración de los mismos y también en su análisis. Para ambas tareas se ha utilizado SPSS, paquete informático sobre el cual los alumnos recibieron un curso del Servicio de Ayuda al Profesorado de la Universidad, y a partir de él, comenzaron a desarrollar una herramienta software que sirviera como puente entre los requerimientos de análisis de los profesores de Marketing, responsables de la base de datos y los resultados obtenidos como respuesta a sus preguntas.

La realización de este proyecto ha requerido de varios estudios previos y el uso de muchos de los conocimientos obtenidos durante el periodo universitario: estadística, programación, arquitectura, bases de datos, etc. El resultado ha sido una aplicación fácil de utilizar por un usuario no experto en ninguna de estas materias, concretamente los profesores del departamento de Comercialización e Investigación de Mercados.

Para facilitar su uso, los autores han añadido un manual desarrollado con TTS Knowledge Force, una herramienta de autor cedida por la casa comercial. El resultado es un Proyecto Fin de Carrera bien estructurado y correcto que además tiene una perspectiva de uso más allá de nuestra Facultad de Informática.

Los resultados más relevantes se han plasmado en dos artículos que han sido aceptados para ser publicados en el Congreso Español de Informática (CEDI, 2010) y en el Congreso de la Sociedad de Estadística e Investigación Operativa (SEIO 2010) y serán publicados en Septiembre.

*Victoria López*

*Directora*



## Resumen

El objeto del proyecto consiste en la realización de un software que permite realizar diversos análisis o estudios sobre encuestas realizadas a consumidores de marcas de distribuidor.

Este proyecto permite la realización de análisis de evolución sobre diversas respuestas de las encuestas con respecto a distintas variables, como pueden ser la edad, status social, género etc. Permite, además, realizar análisis de dependencia entre diversas respuestas de las encuestas usando diferentes niveles de confianza. Por último, el proyecto permite realizar una comparación de consumo de marcas de distribuidor por rangos de edad y olas usando técnicas de Soft-Computing.

El proyecto se basa en un entorno gráfico amigable que permite a los usuarios realizar los diversos estudios de los que dispone el sistema aplicando diversas opciones. También facilita la presentación de los resultados de los estudios mediante la generación de informes detallados. Se pretende que estos estudios sirvan para entender mejor la evolución, tendencia y gustos de los consumidores con respecto a las marcas de distribuidor.

**Palabras clave:** *marcas de distribuidor; MDD; encuestas; SBDataAnalysis; evolución de respuestas; estudios de dependencia; comparación de consumo; enseñanzas; Soft-Computing; MCDM.*

## Abstract

The purpose of the project consists of a software that allows make several store brands surveys analysis.

This project allows the analysis on several survey responses with respect to different variables, such as age, social status, gender and so on. It also allows analysis of dependence between various survey responses using different confidence levels. Finally, the project allows a comparison of store brands consuming by age group and waves using Soft-Computing techniques.

The project is based on a friendly graphical environment that allows users to perform several studies applying different system options. The project facilitates the presentation of the studies results by generating detailed reports. It is intended that these studies will serve to better understand the evolution, trends and tastes of consumers with respect to store brands.

**Keywords:** *store brand; SB; surveys; SBDataAnalysis; evolution of answers; dependence analysis; consume comparison; brand; Soft-Computing; MCDA*





## Introducción

Desde hace más de veinte años, el departamento de *Comercialización e Investigación de Mercados* de la facultad de Ciencias Económicas y Empresariales de la Universidad Complutense de Madrid dedica una parte de sus esfuerzos en el ámbito de la investigación a estudiar el comportamiento de los clientes de los supermercados con respecto a las Marcas de Distribuidor, (MDD).

Para ello anualmente diversos alumnos realizan estudios a pie de campo con entrevistas personales a los consumidores en los que se detalla el nivel de conocimiento de estos últimos sobre estos productos, así como las posibles razones que llevan en última instancia a estas personas a adquirirlos. Paralelamente se realiza otro estudio de toma de datos dentro de los propios establecimientos en los que se controla de forma cuantitativa diversos datos sobre los productos evaluados como son fabricante, precio, altura de colocación o espacio lineal que ocupan.

El principal problema que acarrearán estos estudios para los directores del proyecto de investigación de la facultad es la dificultad para interpretar los resultados obtenidos, derivados sobre todo de cierta caótica organización y de la excesiva complejidad de las soluciones software con las que cuentan. Para ellos la mayoría de los programas de estadísticos con los que cuentan les ofrecen un abanico demasiado amplio y complejo de opciones que no siempre les proporciona el resultado esperado, sino que lo complementan con mucha información poco o nada relevante para su objetivo y que debido a su no especialización en ese campo puede llevar a la interpretación incorrecta de resultados.

El objetivo de desarrollar una herramienta software plenamente adaptada es interesante desde todos los puntos de vista. Las principales metas son saber responder a ciertas preguntas relacionadas con dependencias como puede ser que relación tiene el conocimiento con respecto al consumo. También existe un gran interés en poder establecer un futuro patrón de comportamiento de consumidores según su edad basándose en lo estudiado en años anteriores.

La solución que ofrece este proyecto es desarrollar una interfaz software adaptada plenamente a las necesidades del usuario y que a diferencia de las herramientas estadísticas ya existentes, ofrece una interfaz en la que se pueden realizar los estudios en los que se esté interesado de una forma sencilla e intuitiva. Confiando así en que se pueda progresar de una forma más rápida en las investigaciones y se obtengan resultados que sean susceptibles de ser empleados en acciones futuras.



# 1. Estado del arte

## 1.1. Introducción

Un paquete estadístico es un programa informático que está especialmente diseñado para resolver problemas en el área de la estadística, o bien está programado para resolver problemas de esta área+. Esta es la definición más formal que podemos encontrar buscando por la web. Existen muchos programas que no son especialmente estadísticos pero que pueden hacer algunos cálculos utilizables en estadística aplicada. Estos programas siguen impulsando enormemente la labor de los investigadores que desean utilizar la estadística como apoyo en su trabajo. Los paquetes más sencillos tienen interfaz por ventanas, lo que implica facilidad de uso y aprendizaje pero un mayor grado de complejidad a la hora de hacer cálculos que el programa no tenga predefinidos. Los programas más complejos suelen tener la necesidad de conocer su lenguaje de programación, pero suelen ser mucho más flexibles al poderse incluir en ellos funciones, tests o contrastes que no traen instalados por defecto.

Gracias a la potencia de cálculo de un ordenador, un investigador puede realizar cientos o miles de contrastes de hipótesis en un tiempo muy reducido, al igual que calcular decenas de modelos de regresión en un tiempo muy corto y después quedarse con el más apropiado de ellos. En problemas de investigación de operaciones un programa estadístico es capaz de realizar miles de iteraciones por segundo de un algoritmo en el que una persona tardaría varios minutos en cada una de ellas, lo que proporciona una gran ventaja. Asimismo, es capaz de elegir entre miles de resultados posibles cuál de todos ellos es el óptimo. Básicamente, lo que permiten es resolver problemas de estadística aplicada por fuerza bruta o por probar miles de combinaciones para quedarse finalmente con la que se crea que es la mejor.

En los programas más complejos se necesita tener conocimientos de programación para realizar los cálculos más laboriosos. Por ejemplo en datos tabulados, si se desea realizar una operación dada a una columna concreta, lo más frecuente es que se pueda hacer esto por ventanas, sin embargo, si deseamos hacer esto mismo para todas las columnas de nuestro documento, que pueden ser centenares, es posible que necesitemos programar un bucle en el programa.

Otro inconveniente está en que en estadística a menudo pueden salir resultados contradictorios entre tests distintos, y aunque un programa informático se pueda dar cuenta de ello y avisar de algún modo, tendría que ser el usuario el que decidiera a cuál de los test hacer caso, y esto en ocasiones es complejo, sobre todo si no se está familiarizado a fondo con la estadística o se tiene poca experiencia.

Existen multitud de paquetes informáticos, tanto de software privado como de software libre. Ejemplos son programas como **SAS, SPSS, Excel, Statgraphics y Matlab**. A pesar de que SAS y SPSS son considerados los más potentes, hay muchísimas empresas que utilizan programas mucho menos potentes como Microsoft Excel, debido sobre todo a que su uso es más sencillo, la mayoría de las personas están familiarizadas con él y gran parte de las empresas ya tienen instalado Excel en sus ordenadores, mientras que las licencias de SAS y SPSS cuestan varios miles de euros. Sin embargo, para trabajos más específicos y a nivel de investigación estadística se utilizan siempre paquetes estadísticos especializados ya que suelen tener una capacidad y prestaciones más avanzadas.

## 1.2. Software estadístico

**Matlab** (abreviatura del nombre *MATrix LABoratory*, cuya traducción sería "laboratorio de matrices", <http://www.mathworks.com>) es un programa matemático que tiene un entorno de desarrollo integrado (IDE) con un lenguaje de programación propio (lenguaje M). Algunas de sus prestaciones son la manipulación de matrices, la representación de datos y funciones, la implementación de algoritmos, la creación de interfaces de usuario (GUI) y la comunicación con programas en otros lenguajes y con otros dispositivos hardware. También dispone de dos herramientas, Simulink (plataforma de simulación multidominio) y GUIDE (editor de interfaces de usuario - GUI). Es un software muy extendido para su uso en universidades y centros de investigación y desarrollo y en los últimos años ha aumentado sus opciones, con funciones como crear código VHDL y programar procesadores digitales de señal directamente.

Matlab es un producto propiedad de The Mathworks, y los usuarios están sujetos a un vendor lock-in (se trata de una técnica que utilizan la mayoría de los proveedores para volvernos dependientes a sus soluciones: una vez que nos decidimos por una tecnología dada, los costos reales o percibidos de cambiarse de tecnología son tan grandes que terminamos volviendo a elegir a ese proveedor, aunque su producto no sea el mejor). Recientemente la herramienta adicional llamada MATLAB Builder, bajo la sección de herramientas Application Deployment, se ha introducido para utilizar funciones MATLAB como archivos de biblioteca, pudiendo utilizarse con ambientes de construcción de aplicación .NET o Java aunque esto tiene una desventaja, y es que el computador donde la aplicación tiene que ser utilizada necesita MCR (MATLAB Component Runtime) para que los archivos MATLAB funcionen correctamente. La figura 1 muestra el entorno gráfico de Matlab.

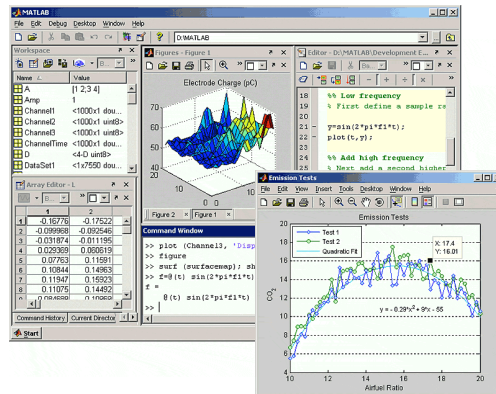


Figura 1. Programa Matlab

**Microsoft Office Excel**, más conocido como [Microsoft Excel](#), es una aplicación para manejar hojas de cálculo. Este programa es desarrollado y distribuido por Microsoft, y es utilizado normalmente en tareas financieras y contables.

En la figura 2 se muestra las tablas y gráficos que se pueden generar en el programa Excel.

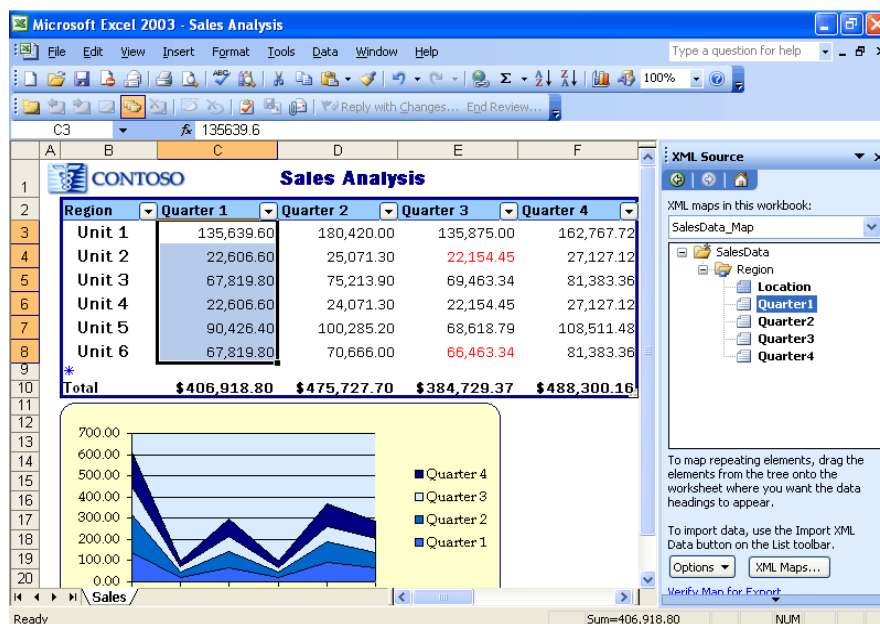


Figura 2. Programa Excel

La versión actual para la plataforma Windows es Excel 12, también denominada Microsoft Office Excel 2007. La versión actual para Mac OS X es la plataforma Microsoft Excel 2008.

Aunque Excel tiene muchas interfaces, ajustadas a las hojas de cálculo electrónico, en esencia sigue siendo la hoja de cálculo como se diseñó en las primeras versiones, VisiCalc: el programa muestra las celdas organizadas en filas y columnas, y cada celda contiene datos o una fórmula, con referencias a otras celdas. Excel introdujo recomputación inteligente de celdas, que hace que las celdas dependientes de otras celdas que han sido modificadas, se actualicen al instante, en vez de como hacían antes programas de hoja de cálculo, que recalculaban la totalidad de los datos todo el tiempo o esperaban algún comando del usuario.

Desde 1993, Excel ha incluido un lenguaje de programación basado en Visual Basic, Visual Basic para Aplicaciones (VBA), que permite automatizar tareas en Excel y proporcionar las funciones definidas por el usuario (UDF) para su uso en las hojas de trabajo y en versiones posteriores, incluye un completo entorno de desarrollo integrado (IDE).

Uno de los mayores problemas de este software de hoja de cálculo, es el que no puede manejar fechas anteriores a 1900, es decir, los campos con un formato de fecha anterior no pueden ser tratados. Esto es así desde las primeras versiones de 16 bits, hasta la actual.

**Statgraphics** (<http://www.statgraphics.net/>) es un software estadístico que realiza funciones matemáticas básicas y avanzadas. El software fue creado en 1980 por el Dr. Neil Polhemus. La versión actual del programa, Statgraphics Centurión XVI, fue lanzado en octubre de 2009.

En la figura 3 se muestra las principales ventanas del programa Stargraphics.

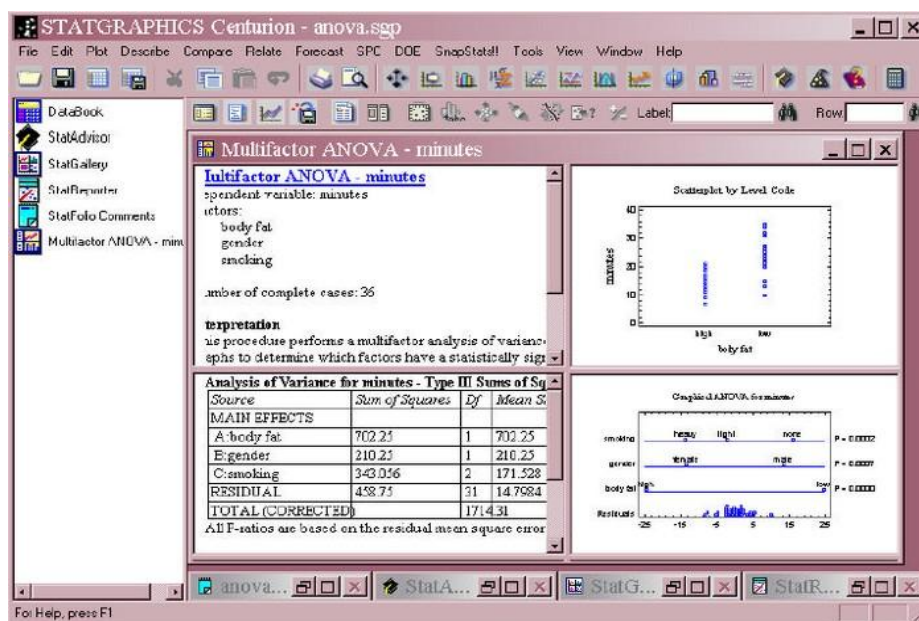


Figura 3. Programa Stargraphics

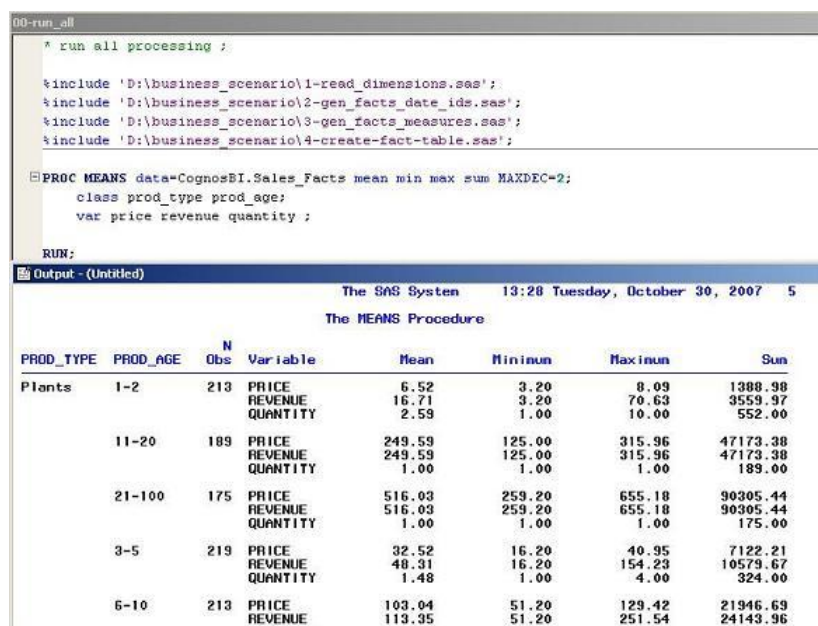
Statgraphics, de la compañía StatPoint Tecnologías Inc, se utiliza con frecuencia en la mejora de procesos con Seis Sigma. El software también se ha utilizado para hacer estudios en el campo de la salud y la nutrición, en diversos estudios relacionados con la estadística.

Statgraphics móvil fue lanzado en el 2006 como el primer programa estadístico de alta capacidad que funciona en ordenadores portátiles (Pocket PC, Pocket PC Phone Edition), o dispositivo compatible con Windows Mobile 5 o Windows Pocket PC 2003, permitiendo una mayor movilidad con la misma potencia de cálculo.

Statgraphics Online es un paquete estadístico que fue lanzado en 2008, y que se ejecuta dentro de un navegador web. Los usuarios pueden importar datos de archivos de texto Excel, u otros formatos, o introducir datos directamente en el editor de datos. Los cálculos se realizan en un servidor web de forma remota y los resultados son devueltos al navegador del usuario como HTML con imágenes de gráficos incrustadas.

**SAS** (<http://www.sas.com/>) posee una manera muy eficiente para transformar los datos estadísticos de diferentes tipos en información, ya que las aplicaciones y sus datos software pueden ser integrados en un solo ambiente software, y así poder entregar a los usuarios la información relevante actualizada.

En la figura 4 se muestra las principales ventanas del programa SAS.



The screenshot shows the SAS software interface. The top window is a code editor titled '00-run\_all' containing SAS code for running all processing, including include statements for various scenarios and a PROC MEANS statement for analyzing sales facts. The bottom window is the 'Output - (Untitled)' window, which displays the results of the PROC MEANS procedure. The output is titled 'The MEANS Procedure' and shows a table of statistics for different product types and ages.

| PROD_TYPE | PROD_AGE | N   | Obs | Variable | Mean   | Minimum | Maximum | Sum      |
|-----------|----------|-----|-----|----------|--------|---------|---------|----------|
| Plants    | 1-2      | 213 |     | PRICE    | 6.52   | 3.20    | 8.09    | 1388.98  |
|           |          |     |     | REVENUE  | 16.71  | 3.20    | 70.63   | 3559.97  |
|           |          |     |     | QUANTITY | 2.59   | 1.00    | 10.00   | 552.00   |
|           | 11-20    | 189 |     | PRICE    | 249.59 | 125.00  | 315.96  | 47173.38 |
|           |          |     |     | REVENUE  | 249.59 | 125.00  | 315.96  | 47173.38 |
|           |          |     |     | QUANTITY | 1.00   | 1.00    | 1.00    | 189.00   |
|           | 21-100   | 175 |     | PRICE    | 516.03 | 259.20  | 655.18  | 90305.44 |
|           |          |     |     | REVENUE  | 516.03 | 259.20  | 655.18  | 90305.44 |
|           |          |     |     | QUANTITY | 1.00   | 1.00    | 1.00    | 175.00   |
|           | 3-5      | 219 |     | PRICE    | 32.52  | 16.20   | 40.95   | 7122.21  |
|           |          |     |     | REVENUE  | 48.31  | 16.20   | 154.23  | 10579.67 |
|           |          |     |     | QUANTITY | 1.48   | 1.00    | 4.00    | 324.00   |
|           | 6-10     | 213 |     | PRICE    | 103.04 | 51.20   | 129.42  | 21946.69 |
|           |          |     |     | REVENUE  | 113.35 | 51.20   | 251.54  | 24143.96 |
|           |          |     |     | QUANTITY | 1.44   | 1.00    | 4.00    | 213.00   |

Figura 4. Programa SAS

El sistema SAS, tal como su nombre indica (Statistical Analysis System), fue desarrollado inicialmente para su uso en tareas de manipulación de datos y análisis estadístico. Es un paquete versátil que se utiliza en diversos ámbitos para el análisis de datos y para la preparación de informes. SAS, es un programa que integra varios módulos interrelacionados, que a la vez están compuestos por procesos simples. Los datos almacenados en formato SAS se pueden usar en cualquier entorno informático (PC, grandes ordenadores, o ) sin ningún cambio, lo que es una gran ventaja en el intercambio de información entre estaciones de trabajo. Para este fin también incluye un módulo específico de comunicaciones llamado SAS/CONNECT.



Inicialmente no se pensó para su uso específico en Economía (como se dice antes, al principio era software puramente estadístico), pero la gran cantidad de funcionalidades ha hecho que se adapte bien para este uso, si bien su uso no está demasiado extendido debido a múltiples razones como el coste de aprendizaje (superior al de la mayoría de paquetes estadísticos), la gran cantidad de documentación que posee (y que complica también su aprendizaje), y el coste de la licencia anual (aunque intentan reducir este coste dividiendo el programa en módulos independientes para que el usuario pueda adquirir sólo aquellos que le interesen)

**SPSS** (<http://www.spss.com/es/>) fue creado como el acrónimo de *Statistical Package for the Social Sciences*, y es un software estadístico muy usado en las empresas que investigan el mercado, así como en ciencias sociales. En la figura 5 se muestra las principales ventanas del programa SPSS.

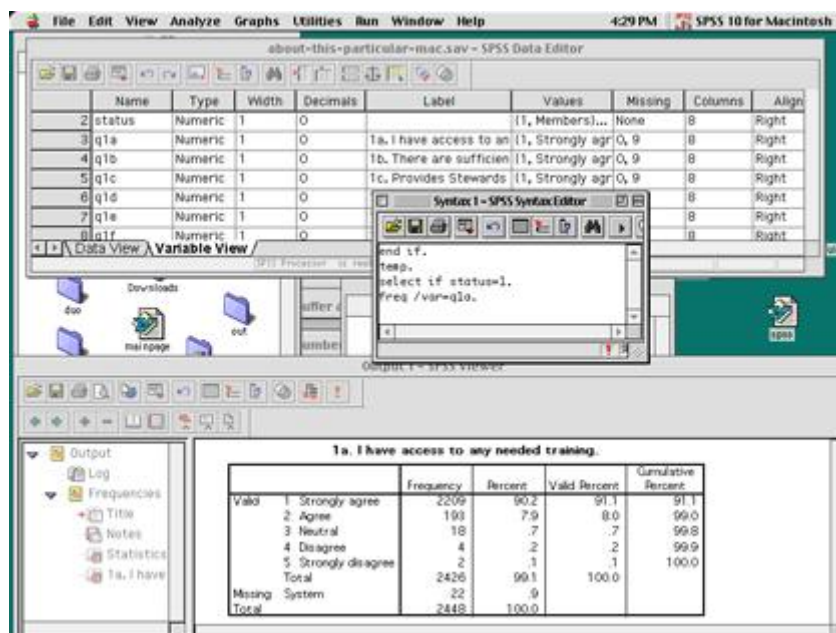


Figura 5. Programa SPSS

Dentro de los programas estadísticos goza de gran popularidad, debido a su capacidad de tratar con bases de datos de gran tamaño, siendo en la versión 12 de SPSS de 250.000 variables y 2 millones de registros, cifra que se va incrementando a medida que surgen nuevas versiones. A esto hay que añadirle la posibilidad de recodificar las variables y registros dependiendo de las necesidades del usuario.

Como muchos de los grandes programas estadísticos, está compuesto por un módulo base que se amplía con módulos anexos, los cuales se venden por separado del módulo principal, y que se va actualizando constantemente para añadirle nuevos procedimientos estadísticos. Actualmente sus principales competidores son SAS y Matlab, aunque también está entrando recientemente en la lucha el software de código abierto y libre, entre los que cabe destacar el lenguaje R, que es un entorno de programación para análisis estadístico y gráfico.



Al igual que para los programas estadísticos más destacados, para poder manejar adecuadamente SPSS hay que leer detenidamente el manual, o realizar algún curso de aprendizaje, ya que los procedimientos pueden tener varias opciones. A pesar de todo esto, se dice que SPSS es el más fácil de aprender para los investigadores, con un manual bastante fácil que explica los mecanismos de las técnicas estadísticas.

El proyecto está orientado a un segmento muy concreto dentro del campo de los estudios estadísticos, y en él se realizan algunos cálculos estadísticos pensados especialmente para los objetivos principales del diseño, no dejando así posibilidad alguna para realizar otras tareas que no sean las planeadas en un principio. Hay que dejar bien claro, que el software que se ha diseñado e implementado no es un paquete estadístico, ni mucho menos, sino que al estar basado en componentes estadísticas implementa funcionalidades que también son resolubles con estos paquetes anteriormente mencionados.

Existen a su vez multitud de programas estadísticos de menor tamaño, tanto gratuitos como de pago, que si que podrían ser, salvando las distancias, comparados a lo que se ha realizado. El problema es, que al igual que el de este proyecto, están orientados a espectros de estudio también muy concretos y la información sobre ellos fuera del círculo para el que están diseñados es escasa.

Un ejemplo en el campo de la detección de epidemias podría ser **EPIDAT** (<http://www.paho.org/Spanish/SHA/epidat.htm>), un proyecto para Microsoft Windows de código abierto iniciado en 1991 por la Dirección General de Salud Pública de la Xunta de Galicia. Como podemos leer en su página web, los objetivos de este software son crear una herramienta de manejo sencillo y de utilidad para el análisis de datos tabulados, así como facilitar la docencia epidemiológica o aumentar la capacidad analítica de los profesionales de la salud. Este software posee posibilidades de estudios demográficos, muestreos o análisis de tipo bayesiano entre otras características.

Como se explica más adelante en esta memoria, una de las componentes principales de del proyecto es el análisis de datos para tratar de definir comportamientos de la población a través de técnicas de Soft-Computing. Pues bien, también existen herramientas software que usan algunas de estas técnicas con fines más lucrativos, como puede ser ganar la lotería. En este campo podemos señalar el **Lotto Wizard** (<http://www.lotto-wizard.net/>), que basándose en combinaciones, estrategias de pesos y detección de combinaciones poco probables, asegura que obtiene mejores resultados que con los típicos métodos de selección de números al azar.

El objetivo de la aplicación va más allá del simple uso de un software estadístico, se trata de permitir hacer un estudio completo sobre las marcas de distribuidor (MDD). Se pretende poder sacar conclusiones concretas sobre una base de datos con mucha información, y por eso no nos sirve un paquete estadístico de uso general, si no que se necesita una aplicación más concreta.

Un ejemplo de estudio de las marcas de distribuidor es el que se comenta en el artículo de la revista *Distribución y Consumo* en la edición de Mayo-Junio del 2003, un artículo del catedrático José Antonio Puelles Pérez y la doctora María Puelles Gallo, y que trata sobre la evolución de las MDD. Dicho artículo, proporciona una explicación del estudio que se ha realizado, con la toma de datos durante varios años, y el intento de estudio de dichos datos mostrándonos por ejemplo, las tendencias de las marcas de distribuidor año tras año y las etapas evolutivas en las diferentes generaciones.

Estos mismos autores publicaron un artículo en el número 100 de la revista *Distribución y Consumo*, en la edición de Julio del 2008, en el que explicaban 100 ideas clave sobre las MDD. Estas ideas recogían datos y reflexiones sobre las iniciativas empresariales más novedosas en el mundo de las Marcas de distribuidor.

Los estudios de marcas de distribuidor no se realizan solo en ambientes académicos. El auge de las marcas de distribuidor en estos tiempos ha hecho que muchas empresas estudien este fenómeno. María Ibañez, Consejero Delegado de la empresa Interbrand, publicó un artículo, en Noviembre del 2008, en la página web de marketing, [www.marketingcomunidad.com](http://www.marketingcomunidad.com), en la que explicaba que en los tiempos de inestabilidad económica el aumento del consumo de las MDD ha sido de un 20% y que son unos grandes aliados para aquellos consumidores que sufren más esta inestabilidad económica.

Habiendo estudiado las diversas opciones que había disponibles para el estudio de bases de datos, y tras sopesar los pros y los contras de cada una de ellas se decidió el uso de la herramienta SPSS ya que era la que parecía más completa y la mejor para tratar el problema.

Una de las principales intenciones era poder unificar toda la información a tratar en un único archivo desde el cual poder hacer todos los estudios para facilitar la manejabilidad del programa, objetivo que se ha conseguido con SPSS, el cual además de poder tratar bases de datos con mucha más información, nos permite la transformación de los archivos en formato Excel (.xls) a .sav (extensión del programa SPSS) de una forma casi automática conservando los nombre de las variables y sus posiciones para su estudio.

Otra de las razones por las que se eligió SPSS es por tener la posibilidad de interactuar con el programa desde aplicaciones exteriores, mediante programación en C#, python ([www.python.org](http://www.python.org)) o R como se explica más adelante, y esto es exactamente lo que se pretende con el proyecto.

Se piensa que esta herramienta (la del proyecto) es mejor que las anteriores dedicadas a los mismos fines ya que para toda herramienta que sirve para estos propósitos, es necesario hacer un curso de iniciación a la herramienta, el cual es costoso, largo y complicado para el usuario, mientras que con este software, contando con la misma potencia que el software SPSS ya que lo ejecutamos en modo backend, no se necesitará ningún tipo de curso de iniciación ya que es sumamente intuitiva, y se cuenta con la pestaña help que muestra información sobre las variables de la base de datos, y resúmenes sobre las distintas pestañas que hay.

## 2. Conceptos teóricos previos

### 2.1. Marcas de Distribuidor

Las marcas de distribuidor (MDD), también conocidas como marcas blancas o marcas genéricas, son las marcas pertenecientes a una cadena de distribución (generalmente supermercados o hipermercados, pero también de gran distribución especializada) con la que se venden productos de distintos fabricantes. Este tipo de marcas han considerado tradicionalmente la estrategia *Business to Consumer (B2C)*, estrategia que desarrollan las empresas comerciales para llegar directamente al cliente o usuario final.

Los primeros envases de estos artículos eran blancos y bastante asépticos, de ahí el nombre de marca blanca. Trataban de captar al comprador innovador haciendo ver que ha hecho la mejor elección, y no está pagando de más por un artículo básico.

Los tres mayores campos dentro de las MDD son los productos de alimentación, los productos de limpieza del hogar y los productos de higiene personal, productos de compra frecuente por parte del consumidor. Las MDD aparecen y se consolidan para artículos de compra de menor exigencia por parte del comprador o usuario final. Las marcas de distribuidor, como norma general, son más baratas que las marcas de fabricante. Esto es debido fundamentalmente al ahorro producido por la reducción de los costes de publicidad, aunque en estos momentos ha aumentado la su promoción. En muchos casos, son productos idénticos a los comercializados por las marcas líderes, o marcas de fabricante (MF), pero a un coste inferior. El fabricante de las marcas de distribuidor, al ser una empresa de distribución, tiene asegurada la implantación de su producto en un mercado concreto y amplio, sus propios puntos de venta. Las marcas de distribuidor producen una percepción de los consumidores de que el control de calidad pueda ser menos riguroso para ahorrar costes; o que se intente ahorrar en la materia prima. Con los productos de marcas de distribuidor se pierde la conexión entre el fabricante y el cliente, esto puede ser importante para artículos con servicio postventa, como electrodomésticos, aparatos de audio o video etc.

En los últimos años algunas conocidas empresas (por ejemplo Kellogg o Danone) están anunciando notoriamente que no fabrican productos para MDD. Algunas lo hacían en el pasado (como en el caso de Danone) pero decidieron cambiar de política y dejar de hacerlo, apostando por sus propias marcas. El auge de las MDD ha perjudicado a algunas marcas de fabricante, lo que ha hecho que inviertan menos en publicidad, y esto a su vez ha recudido los ingresos de algunos medios de comunicación. Debido a ello algunos medios se han posicionado recientemente a favor de las marcas de fabricante. En la figura 6 se puede observar dos bricks de leche, uno de marca de distribuidor y otro de marca de fabricante, y aparentemente no tienen ninguna diferencia.



Figura 6. Un Brick de leche de MDD y otro brick de MF

Las marcas de distribuidor han conseguido situarse en el mercado como marcas con nombre propio, tratando de equipararse a las marcas de fabricante y competir con él. Para ello han creado especializaciones. A continuación se muestra un listado con algunas de estas enseñas en el mercado español:

- **AhorraMas:** Alipende (alimentación), Meque (mascotas), Lanta (droguería), Bodyplus (higiene personal).
- **Alcampo:** Primer Precio (la de menor precio-calidad), Auchan (Alimentación), Basic Line (imagen, sonido, electrodomésticos), Rik&Rok (infantil), Cup's (deportes), Genium (bricolaje, automóvil), In Extenso (textil).
- **Carrefour:** Carrefour (en la mayoría de departamentos), 1 (en la mayoría de departamentos, se supone que la de menor precio-calidad), Tex (textil), Basic (textil), First Line (electrodomésticos), Blue Sky (electrodomésticos), Top Bike (deportes).
- **Dia:** Dia.
- **Dinosol Supermercados, S.L. (supersol, Hiperdino):** superSol.
- **D&S Líder:** Líder, Acuenta (Alimentos), Groven (Electro).
- **E.Leclerc:** Marca Guía, Eco+ (se supone que la de menor precio y quizá calidad).
- **El Árbol:** Super.
- **El Corte Inglés:** El Corte Inglés (alimentos), HiperCor (alimentos), Aliada (alimentos), Boomerang (deportes), Emidio Tucci (ropa) fórmula joven (ropa), Gloria Ortiz (ropa), digrato (electrodomésticos), inves (aparatos electrónicos).
- **Eroski-Caprabo:** Consumer, Eroski, Consum, Erokson (electro, no se usa actualmente), Ecron (electrodomésticos, imagen, sonido), Romester (deportes), Visto Bueno (textil).

- **Jumbo:** Jumbo (Alimentos), GlobalHome (Electro).
- **Mercadona:** Hacendado, Bosque Verde (droguería), Deliplus (droguería).
- **MásyMás:** Alteza (pastas), es Gallo.

## 2.2. Etapas evolutivas en las MDD

Las MDD han ido evolucionando desde su creación a mediados del s.XX hasta nuestros días. Cada evolución (o generación) ha sido motivada por los diferentes cambios que ha habido en la sociedad durante este siglo y medio. Detrás de cada una de estas generaciones ha habido un gran estudio por parte de los fabricantes para entender cuál ha sido la tendencia de consumo y adecuar sus productos a esas tendencias.

- **Primera Generación. El precio como única referencia**

Comienzan en EEUU en la década de los sesenta con la aparición de productos genéricos, que son productos con etiquetas muy simples y sin nombre haciendo tan sólo referencia al producto; sus precios bajos estaban amparados en la buena imagen del establecimiento que los comercializa.

El fenómeno se traslada a España con los productos *Simago*, que aparecen en la segunda parte de la década de los sesenta. Estos productos Simago estaban contenidos en unos envases muy sencillos, que sin más indicaban el producto contenido y el logotipo de Simago.

- **Segunda generación. Precio bajo e intento de aproximarse a la calidad del líder**

*Carrefour* en 1976 introduce una nueva forma de competencia consistente en lanzar cincuenta productos libres o sin marca de fabricante, con el fin de diferenciar el producto al incorporarle otra marca, la propia, ofreciendo precios competitivos sin disminuir la calidad.

Las MDD nacen por la necesidad del distribuidor de buscar su propia identidad y de crear una imagen propia para sus establecimientos, suponen incluso un medio de comunicación para los distribuidores. Se trata de una herramienta con doble finalidad para el distribuidor: por un lado traslada su imagen al mercado y, por otro, es de gran utilidad por la creación de la lealtad de la clientela hacia su propio establecimiento. Por tanto, sirven como medio de comunicación y de diferenciación de los detallistas.

- **Tercera generación. Tiendas de descuento. MDD como soporte del establecimiento**

Surgen en España, a principios de la década de los 80 las tiendas de descuento, de la mano de una cadena, *Dia %* (Distribuidora Internacional de Alimentación) que cosecha éxitos interrumpidos desde su constitución hasta hoy. *Dia %* implanta en España una nueva fórmula, el descuento, con supermercados situados inicialmente en el casco urbano de Madrid, con surtido reducido, con productos de buena calidad, un gran número de establecimientos y apoyados en las MDD como soporte del establecimiento.

- **Cuarta generación. Igual calidad que el líder pero con precios más bajos**

También a principios de la década de los 80 se busca de forma directa y como objetivo la comparación con el líder en cada categoría y se intenta transmitir que la diferencia con el líder no está en la calidad, sino sólo en el precio.

Para intentar transmitir este mensaje los distribuidores encargan la fabricación de sus MDD a fabricante de marcas líderes o fabricantes exclusivos de MDD pero con tecnología también de vanguardia, pidiendo a unos y otros que las MDD que les fabriquen se parezcan lo máximo posible a las marcas líderes.

- **Quinta generación. Primer intento de la segunda calidad de MDD**

A mediados de los años ochenta se hace un intento de que un mismo distribuidor ofreciese a la vez en sus tiendas dos, o incluso más, calidades diferentes de MDD. Este primer intento en España se produce por la organización francesa *Promodes*, a través de la enseña *Simago*, y con su marca ~~Nofrills~~ (sin adornos) que tuvo escaso éxito.

- **Sexta generación. MDD en el Hard Discount (descuento duro alemán) como soporte de sus establecimientos**

El descuento duro (existente en los Hard Discount) fija unos precios aún más bajos que los de descuento (como el ya comentado de la cadena *Dia %*), consistente en surtido reducido, estricto control de costes, bajos márgenes de beneficio y apoyándose aún más en las MDD, más simples y con una menor imagen de calidad.

El negocio del descuento duro alemán al que estamos haciendo referencia se apoya de una forma esencial en sus MDD que no coinciden con la enseña. Las MDD empleadas tienden a adaptarse al país donde operan. Por otra parte en estos establecimientos la presencia de las marcas de fabricante en general y de marcas líderes de fabricantes en particular, es meramente testimonial. Este tipo de negocio comenzó a implantarse en España a mediados de los ochenta.

- **Séptima generación. Singularidad de algunos productos cuyas MDD no pueden coincidir con la enseña (vinos, cavas, bebidas de alta graduación etílica etc.)**

A principios de la década de los 90 hay algunos productos que por su relevancia del entorno social en el momento en que se consumen no pueden salir al mercado con unas MDD coincidentes con la enseña.

Por ejemplo, invitar a unos amigos y ofrecerles un buen vino con denominación de origen Rioja, Rivera de Duero con marca ~~Δ~~ia %qo ~~Δ~~El Corte Inglésqes impensable.

- **Octava generación. Marca de distribuidor coincidente con la enseña, de calidad percibida superior a la del líder.**

Se acepta que la calidad percibida de un producto líder justifica un precio superior para este tipo de productos. En el fenómeno de las MDD el diferencial del precio entre la marca de fabricante y el producto con marca de distribuidor mucho tiene que ver con la comparación de la imagen de calidad de quien elabora el concreto producto líder y el distribuidor que lo vende en sus establecimientos.

En la mayoría de los casos el diferencial de precios entre marca de distribuidor y marca de fabricante normalmente es negativa, es decir, el precio de los productos de marca de fabricante es mayor que los precios de los productos de marca de distribuidor. Hay un caso en el que este diferencial es positivo, en el distribuidor ~~Δ~~El Corte Inglésq que tiene una muy buena imagen de calidad, podemos ver algunos muy determinados productos en los que los precios del producto con marca de distribuidor de El Corte Inglés son superiores a los que de ese mismo producto de un líder, que es su fabricante. Este sería el caso, por ejemplo, de la leche entera en envases de un litro en brick, en el que la marca de distribuidor tiene un precio más elevado que la marca de fabricante que le suministra el producto (Central Lechera Asturiana).

- **Novena generación. MDD no coincidente con la enseña, con calidad percibida superior a la del líder**

Imitando en cierta medida el caso descrito de El Corte Inglés, algunos distribuidores a finales de los noventa lanzaron productos de superior calidad en algunas referencias (y cuando no existe un líder tan indiscutible como Danone o Coca-Cola).

Este es el caso de los productos de Carrefour con MDD %De nuestra tierra+ que como productos selectos y especiales se venden a precios superiores a los correspondientes con marcas de fabricante.

- **Décima generación. Concentración de MDD. Proyección transnacional. Aparición de fabricantes especialista que no ofrecen sus marcas en el mercado**

Las MDD en su crecimiento y en la búsqueda de economías de escala y de unificación en la comunicación llegó un momento, a principios de este siglo en el que las MDD tenían el mismo nombre, generalmente coincidente con la enseña en los distintos países en que el distribuidor se iba proyectando. Surgen especialistas que crean con el único objetivo de fabricar MDD, sin estar presentes en el mercado, para ofrecer sus marcas.

- **Undécima generación. Doble comparación, con el líder y con primer precio**

Las marcas denominadas de primer precio son aquellas marcas identificadas con el precio más bajo del lineal. Los distribuidores solían utilizar las marcas del primer precio como una segunda marca de distribuidor no identificada claramente con la enseña, que le permitía subir el nivel de calidad y de margen de las MDD y competir a dos bandas con las marcas de fabricantes. Esta tendencia empezó a darse a partir del año 2004.

- **Duodécima generación. Últimas tendencias**

- **Segunda concentración de MDD. Fabricación y comercialización a nivel multinacional**

No faltan quienes opinan que las empresas que tienen dimensión nacional en un mundo tan competitivo como el actual, tienen los días contados y que, antes o después, se integrarán de una forma u otra en organizaciones concebidas, bien a nivel multinacional, bien en empresas que se proyecten al menos a nivel de la UE o a un nivel multinacional ya que en otro caso dejarán de ser competitivas.

Los grandes distribuidores están concentrando sus MDD en pocas marcas, generalmente coincidentes con la enseña y que se venden en todos los países en los que están presentes. Concentran también la producción de las mismas, aunque lógicamente de forma distinta que los grandes industriales, solicitando ofertas a industriales con capacidad para fabricar en uno u otro país MDD, que acaban suministrando a los establecimientos que esos grandes distribuidores tienen en varios países.

- **Segunda versión de la segunda calidad de MDD. Proyección plurinacional.**

Las grandes empresas de MDD sacan al mercado su segunda calidad y ahora sí parece que son acogidas de mejor manera por los compradores.



- **Actualmente**

Las MDD están creando nuevas iniciativas hacia el precio bajo como camino para fidelizar a dos colectivos diferentes pero relevantes ambos:

- Las poblaciones marginales obligadas a comprar el precio mas bajo.
- La población reflexiva y racional en sus comportamientos.

## 2.3. Análisis Estadístico de MDD

En esta sección se explican los elementos estadísticos necesarios para el análisis de MDD. Comienza la explicación con la recogida de datos mediante encuestación, y a continuación se explican los estadísticos y técnicas estadísticas más utilizadas para el análisis de mercados de las MDD.

Una encuesta es un conjunto de preguntas normalizadas dirigidas a una muestra representativa de la población o instituciones, con el fin de conocer estados de opinión o hechos específicos. El investigador debe seleccionar las preguntas más convenientes, de acuerdo con la naturaleza de la investigación y, sobre todo, considerando el nivel de educación de las personas que se van a responder el cuestionario. Las encuestas tienen por objetivo obtener información estadística indefinida, mientras que los censos y registros vitales de población son de mayor alcance y extensión. Este tipo de estadísticas pocas veces otorga, en forma clara y precisa, la verdadera información que se requiere, de ahí que sea necesario realizar encuestas a esa población en estudio, para obtener los datos que se necesitan para un buen análisis. Este tipo de encuesta abarca generalmente el Universo de los individuos en cuestión.

Otro tipo de encuestas es *Encuestas por Muestreo* en donde se elige una parte de la población que se estima representativa de la población total. Debe tener un diseño muestral, necesariamente debe tener un marco de donde extraerla y ese marco lo constituye el censo de población. La encuesta (muestra o total), es una investigación estadística en que la información se obtiene de una parte representativa de las unidades de información o de todas las unidades seleccionadas que componen el universo a investigar. La información se obtiene tal como se necesita para fines estadístico-demográficos.

Una forma reducida de una encuesta por muestreo es un "*sondeo de opinión*", esta forma de encuesta es similar a un muestreo, pero se caracteriza porque la muestra de la población elegida no es suficiente para que los resultados puedan aportar un informe confiable. Se utiliza solo para recolectar algunos datos sobre lo que piensa un número de individuos de un determinado grupo sobre un determinado tema.

Actualmente, existen sistemas de gestión de encuestas en Internet, que están acercando su utilización a investigadores que, hasta el momento, no tenían acceso a los medios necesarios para ejecutarlas.

Algunas aplicaciones de las encuestas se muestran a continuación:

- Medir las relaciones entre variables demográficas, económicas y sociales.
- Evaluar las estadísticas demográficas como errores, omisiones e inexactitudes.
- Conocer profundamente patrones de las variables demográficas y sus factores asociados como fecundidad y migraciones determinantes.
- Otorga información suplementaria en relación a la otorgada por los Censos.
- Evaluar periódicamente los resultados de un programa en ejecución.
- Probar la eficiencia de un método antes de aplicarlo al total de la población.
- Saber la opinión o acción del público acerca de un determinado tema.

Algunas de las ventajas de una encuesta son las siguientes:

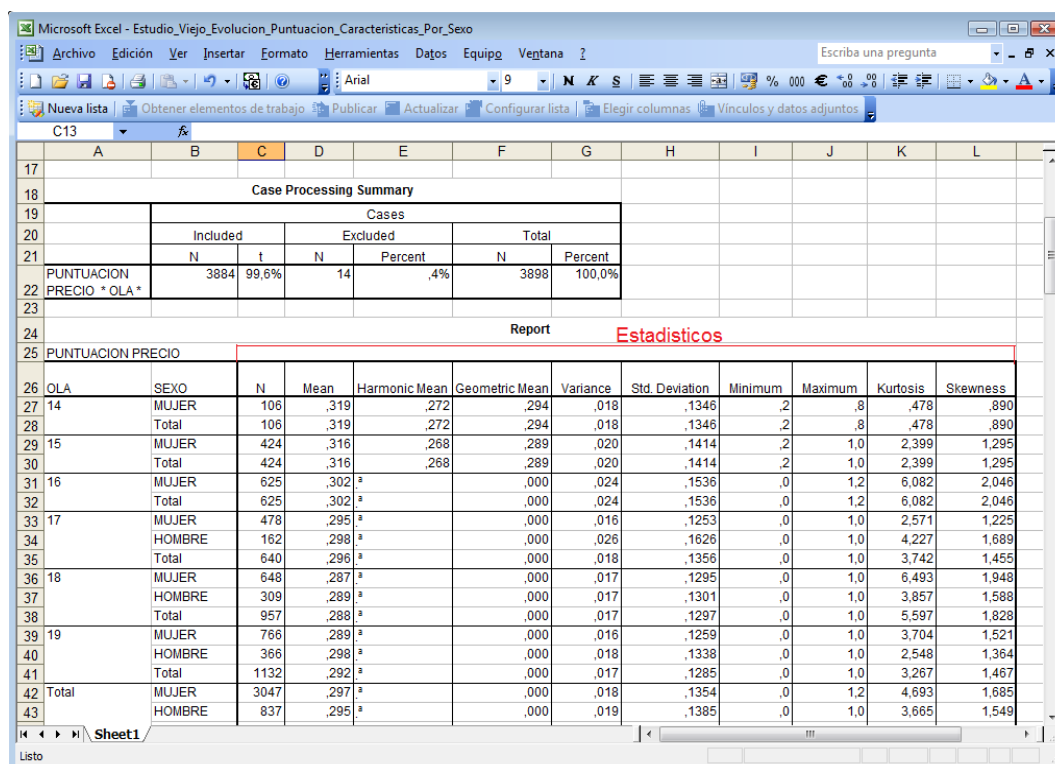
- Bajo costo
- Información más exacta (mejor calidad) que la del Censo debido al menor número de empadronadores permite capacitarlos mejor y más selectivamente.
- Es posible introducir métodos científicos objetivos de medición para corregir errores.
- Mayor rapidez en la obtención de resultados.
- Técnica más utilizada y que permite obtener información de casi cualquier tipo de población.
- Permite obtener información sobre hechos pasados de los encuestados.
- Gran capacidad para estandarizar datos, lo que permite su tratamiento informático y el análisis estadístico.
- Relativamente barata para la información que se obtiene con ello.
- Te ayuda a conocer lo que quisieras conocer de la persona o personas encuestadas

Algunas desventajas de una encuesta pueden ser:

- El planeamiento y ejecución de la investigación suele ser más complejo que si se realizara por censo.
- Requiere para su diseño de profesionales con buenos conocimientos de teoría y habilidad en su aplicación.
- Un mal entrenamiento al personal puede inferir en la encuesta.
- Los encuestados no sean sinceros en la contestación de las preguntas.
- Gastos excesivos al hacer una mala creación de la encuesta.

En el anexo 2a se muestra una guía de investigación que el departamento de comercialización e investigación del mercado entrega a los encuestadores y en el que se explica las instrucciones tanto de la encuesta a consumidores como de la encuesta de producto. En el anexo 2b se muestra una encuesta de consumidores, [Información facilitada por el departamento de comercialización e investigación del mercado de la Universidad Complutense de Madrid]

Un estadístico es una función que transforma los datos de un muestreo en un valor. Dicho valor se utiliza como estimador de algún dato de interés. A continuación se definen los diferentes estadísticos que se usan en los diversos estudios que la aplicación realiza. Estos estadísticos permiten analizar distintas características de los resultados, obteniendo una visión del estudio mucho más global al usuario, que podrá obtener mejores conclusiones. En la figura 7 se muestra un archivo Excel con los resultados de un análisis, con los estadísticos que se calculan



| Case Processing Summary   |  |          |       |          |         |       |         |  |  |  |  |
|---------------------------|--|----------|-------|----------|---------|-------|---------|--|--|--|--|
|                           |  | Included |       | Excluded |         | Total |         |  |  |  |  |
|                           |  | N        | t     | N        | Percent | N     | Percent |  |  |  |  |
| PUNTUACION PRECIO * OLA * |  | 3884     | 99,6% | 14       | ,4%     | 3898  | 100,0%  |  |  |  |  |

| Report Estadísticos |        |      |                   |               |                |          |                |         |         |          |          |
|---------------------|--------|------|-------------------|---------------|----------------|----------|----------------|---------|---------|----------|----------|
| OLA                 | SEXO   | N    | Mean              | Harmonic Mean | Geometric Mean | Variance | Std. Deviation | Minimum | Maximum | Kurtosis | Skewness |
| 14                  | MUJER  | 106  | ,319              | ,272          | ,294           | ,018     | ,1346          | ,2      | ,8      | ,478     | ,890     |
| 28                  | Total  | 106  | ,319              | ,272          | ,294           | ,018     | ,1346          | ,2      | ,8      | ,478     | ,890     |
| 29                  | MUJER  | 424  | ,316              | ,268          | ,289           | ,020     | ,1414          | ,2      | 1,0     | 2,399    | 1,295    |
| 30                  | Total  | 424  | ,316              | ,268          | ,289           | ,020     | ,1414          | ,2      | 1,0     | 2,399    | 1,295    |
| 31                  | MUJER  | 625  | ,302 <sup>a</sup> |               | ,000           | ,024     | ,1536          | ,0      | 1,2     | 6,082    | 2,046    |
| 32                  | Total  | 625  | ,302 <sup>a</sup> |               | ,000           | ,024     | ,1536          | ,0      | 1,2     | 6,082    | 2,046    |
| 33                  | MUJER  | 478  | ,295 <sup>a</sup> |               | ,000           | ,016     | ,1253          | ,0      | 1,0     | 2,571    | 1,225    |
| 34                  | HOMBRE | 162  | ,298 <sup>a</sup> |               | ,000           | ,026     | ,1626          | ,0      | 1,0     | 4,227    | 1,689    |
| 35                  | Total  | 640  | ,296 <sup>a</sup> |               | ,000           | ,018     | ,1356          | ,0      | 1,0     | 3,742    | 1,455    |
| 36                  | MUJER  | 648  | ,287 <sup>a</sup> |               | ,000           | ,017     | ,1295          | ,0      | 1,0     | 6,493    | 1,948    |
| 37                  | HOMBRE | 309  | ,289 <sup>a</sup> |               | ,000           | ,017     | ,1301          | ,0      | 1,0     | 3,857    | 1,588    |
| 38                  | Total  | 957  | ,288 <sup>a</sup> |               | ,000           | ,017     | ,1297          | ,0      | 1,0     | 5,597    | 1,828    |
| 39                  | MUJER  | 766  | ,289 <sup>a</sup> |               | ,000           | ,016     | ,1259          | ,0      | 1,0     | 3,704    | 1,521    |
| 40                  | HOMBRE | 366  | ,298 <sup>a</sup> |               | ,000           | ,018     | ,1338          | ,0      | 1,0     | 2,548    | 1,364    |
| 41                  | Total  | 1132 | ,292 <sup>a</sup> |               | ,000           | ,017     | ,1285          | ,0      | 1,0     | 3,267    | 1,467    |
| 42                  | Total  | 3047 | ,297 <sup>a</sup> |               | ,000           | ,018     | ,1354          | ,0      | 1,2     | 4,693    | 1,685    |
| 43                  | HOMBRE | 837  | ,295 <sup>a</sup> |               | ,000           | ,019     | ,1385          | ,0      | 1,0     | 3,665    | 1,549    |

Figura 7. Los diferentes estadísticos que se usan en un estudio

La **media aritmética** de un conjunto finito de datos es igual a la suma de todos sus valores dividida por el número de sumandos. Cuando el conjunto es una muestra aleatoria recibe el nombre de media muestral siendo uno de los principales estadísticos muestrales.

Expresada de forma más intuitiva, podemos decir que la media aritmética es la cantidad total de la variable distribuida a partes iguales entre cada observación.

Por ejemplo, si en una habitación hay tres personas, la media de dinero que tienen en sus bolsillos sería el resultado de tomar todo el dinero de los tres y dividirlo a partes iguales entre cada uno de ellos. Es decir, la media es una forma de resumir la información de una distribución (dinero en el bolsillo) suponiendo que cada observación (persona) tendría la misma cantidad de la variable.

También la media aritmética puede ser denominada como centro de gravedad de una distribución, el cual no está necesariamente en la mitad.

Una de las limitaciones de la media es que se ve afectada por valores extremos; valores muy altos tienden a aumentarla mientras que valores muy bajos tienden a reducirla, lo que implica que puede dejar de ser representativa de la población.

Matemáticamente, dados los  $n$  números  $a_1, a_2, \dots, a_n$ , la media aritmética se define como:

$$\bar{x} = \frac{\sum_{i=1}^n a_i}{n} = \frac{a_1 + \dots + a_n}{n}$$

La **media armónica**, denotada  $H$ , de una cantidad finita de números es igual al recíproco, o inverso, de la media aritmética de los recíprocos de dichos números

Dados los números  $a_1, a_2, \dots, a_n$ , la media armónica será igual a:

$$H = \frac{n}{\sum_{i=1}^n \frac{1}{a_i}} = \frac{n}{\left(\frac{1}{a_1} + \dots + \frac{1}{a_n}\right)}$$

La media armónica resulta poco influida por la existencia de determinados valores mucho más grandes que el conjunto de los otros, siendo en cambio sensible a valores mucho más pequeños que el conjunto. La media armónica no está definida en el caso de la existencia en el conjunto de valores nulos. La media armónica se suele utilizar para promediar velocidades, tiempos, rendimiento, etc. La media armónica considera todos los valores de la distribución y, en ciertos casos es más representativa que la media aritmética; sin embargo, está muy influenciado por valores pequeños y no se puede determinar en las distribuciones con algunos valores igual a cero.

La **media geométrica** de una cantidad arbitraria de números (digamos  $n$  números) es la raíz  $n$ -ésima del producto de todos los números.

$$\bar{x} = \sqrt[n]{\prod_{i=1}^n x_i} = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n}$$

Por ejemplo, la media geométrica de 2 y 18 es  $\sqrt[2]{2 \cdot 18} = \sqrt{36} = 6$

Otro ejemplo, la media de 1, 3 y 9 sería  $\sqrt[3]{1 \cdot 3 \cdot 9} = \sqrt[3]{27} = 3$

La pregunta que contesta la media geométrica es: si todas las cantidades  $n$  fueran iguales, ¿cuál sería esa  $n$  para que el producto fuera el mismo? El logaritmo de la media geométrica es igual a la media aritmética de los logaritmos de los valores de la variable. La media geométrica considera todos los valores y es menos sensible que la media aritmética a los valores extremos, sin embargo, su cálculo es más difícil y es menos intuitivo que la media aritmética.

Solo es relevante la media geométrica si todos los números son positivos. Como hemos visto, si uno de ellos es 0, entonces el resultado es 0. Si hubiera un número negativo (o una cantidad impar de ellos) entonces la media geométrica sería o bien negativa, o bien inexistente en los números reales. En muchas ocasiones se utiliza su transformación en el manejo estadístico de variables con distribución no normal. La media geométrica es relevante cuando varias cantidades son multiplicadas para producir un total.

En teoría de probabilidad, la **varianza** de una variable aleatoria es una medida de su dispersión definida como la esperanza del cuadrado de la desviación de dicha variable respecto a su media.

Está medida en unidades distintas de las de la variable. Por ejemplo, si la variable mide una distancia en metros, la varianza se expresa en metros al cuadrado. La desviación típica, la raíz cuadrada de la varianza, es una medida de dispersión alternativa expresada en las mismas unidades.

Hay que tener en cuenta que la varianza puede verse muy influida por los outliers (dato que numéricamente está muy distanciado de los demás datos) y se desaconseja su uso cuando las distribuciones de las variables aleatorias tienen colas pesadas. En tales casos se recomienda el uso de otras medidas de dispersión más robustas.

El término *varianza* fue acuñado por Ronald Fisher en un artículo de 1918 titulado *The Correlation Between Relatives on the Supposition of Mendelian Inheritance*.

Se define la varianza  $\text{Var}(X)$  (también representada como  $\sigma_X^2$ , simplemente  $\sigma^2$ ) de  $X$ , una variable aleatoria que tiene como media  $\mu = E(X)$ , como:

$$\text{Var}(X) = E[(X - \mu)^2]$$

Desarrollando la definición anterior, se obtiene la siguiente definición alternativa (y equivalente):

$$\begin{aligned}\text{Var}(X) &= E[(X - \mu)^2] \\ &= E[X^2 - 2X\mu + \mu^2] \\ &= E(X^2) - 2\mu E(X) + \mu^2 \\ &= E(X^2) - 2\mu^2 + \mu^2 \\ &= E(X^2) - \mu^2\end{aligned}$$

Algunas propiedades de la varianza son:

- $V(X) \geq 0$
- $V(aX + b) = a^2 V(X)$  siendo  $a$  y  $b$  números reales cualesquiera. De esta propiedad se deduce que la varianza de una constante es cero, es decir,  $V(b) = 0$ .
- $V(X + Y) = V(X) + V(Y) + 2Cov(X, Y)$ , donde  $Cov(X, Y)$  es la covarianza de  $X$  e  $Y$  (estadístico que nos indica si existe dependencia directa( $<0$ ), inversa( $>0$ ) o si no existe dependencia ( $=0$ ) entre  $X$  e  $Y$ ).
- $V(X - Y) = V(X) + V(Y) - 2Cov(X, Y)$ , donde  $Cov(X, Y)$  es la covarianza de  $X$  e  $Y$ .

La **desviación estándar** o **desviación típica** ( ) es una medida de centralización o dispersión para variables de razón (ratio o cociente) y de intervalo, de gran utilidad en la estadística descriptiva.

Se define como la raíz cuadrada de la varianza, más interesante al medir en las mismas unidades que la media ya que la varianza mide en  $n^2$ . Junto con este valor, la desviación típica es una medida (cuadrática) que informa de la media de distancias que tienen los datos respecto de su media aritmética, expresada en las mismas unidades que la variable.

Para conocer con detalle un conjunto de datos, no basta con conocer las medidas de tendencia central, sino que necesitamos conocer también la desviación que representan los datos en su distribución respecto de la media aritmética de dicha distribución, con objeto de tener una visión de los mismos más acorde con la realidad a la hora de describirlos e interpretarlos para la toma de decisiones.

La varianza representa la media aritmética de las desviaciones con respecto a la media que son elevadas al cuadrado. Si atendemos a la colección completa de datos (la población en su totalidad) obtenemos la varianza poblacional; y si por el contrario prestamos atención sólo a una muestra de la población, obtenemos en su lugar la varianza muestral. Las expresiones de estas medidas son las que aparecen a continuación.

Expresión de la desviación estándar:

$$\sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^N (X_i - \mu)^2}{N}}$$

El término *desviación estándar* fue incorporado a la estadística por Karl Pearson en 1894. Por la formulación de la varianza podemos pasar a obtener la desviación estándar, tomando la raíz cuadrada positiva de la varianza. Así, si efectuamos la raíz de la varianza muestral, obtenemos la desviación típica muestral; y si por el contrario, efectuamos la raíz sobre la varianza poblacional, obtendremos la desviación típica poblacional.

La desviación estándar es una medida del grado de dispersión de los datos con respecto al valor promedio. Dicho de otra manera, la desviación estándar es simplemente el "promedio" o variación esperada con respecto a la media aritmética.

Por ejemplo, las tres muestras (0, 0, 14, 14), (0, 6, 8, 14) y (6, 6, 8, 8) cada una tiene una media de 7. Sus desviaciones estándar son 8.08, 5.77 y 1.15, respectivamente. La tercera muestra tiene una desviación mucho menor que las otras dos porque sus valores están más cerca de 7.

La desviación estándar puede ser interpretada como una medida de incertidumbre. La desviación estándar de un grupo repetido de medidas nos da la precisión de éstas. Cuando se va a determinar si un grupo de medidas está de acuerdo con el modelo teórico, la desviación estándar de esas medidas es de vital importancia: si la media de las medidas está demasiado alejada de la predicción (con la distancia medida en desviaciones estándar), entonces consideramos que las medidas contradicen la teoría. Esto es coherente, ya que las mediciones caen fuera del rango de valores en el cual sería razonable esperar que ocurrieran si el modelo teórico fuera correcto. La desviación estándar es uno de tres parámetros de ubicación central; muestra la agrupación de los datos alrededor de un valor central (la media o promedio).

Entre las técnicas estadísticas más utilizadas en este trabajo se encuentran los contrastes de hipótesis. Un **contraste de hipótesis** (también denominado test de hipótesis o prueba de significación) es una metodología de inferencia estadística para juzgar si una propiedad que se supone cumple una población estadística es compatible con lo observado en una muestra de dicha población. Fue iniciada por Ronald Fisher y fundamentada posteriormente por Jerzy Neyman y Karl Pearson.

Mediante esta teoría, se aborda el problema estadístico considerando una hipótesis determinada  $H_0$ , o hipótesis nula, y una hipótesis alternativa  $H_1$ , y se intenta dirimir cuál de las dos es la hipótesis verdadera, tras aplicar el problema estadístico a un cierto número de experimentos.

Está fuertemente asociada a los considerados errores de tipo I y II en estadística, que definen respectivamente, la posibilidad de tomar un suceso verdadero como falso, o uno falso como verdadero.

Existen diversos métodos para desarrollar dicho test, minimizando los errores de tipo I y II, y hallando por tanto con una determinada potencia, la hipótesis con mayor probabilidad de ser correcta. Los tipos más importantes son los test centrados, de hipótesis y alternativa simple, aleatorizados,... Dentro de los test no paramétricos, el más extendido es probablemente el test de la U de Mann-Whitney.

La aplicación de cálculos probabilísticos permite determinar a partir de qué valor debemos rechazar la hipótesis garantizando que la probabilidad de cometer un error es un valor conocido a priori. Las hipótesis pueden clasificarse en dos grupos, según: Especifiquen un valor concreto o un intervalo para los parámetros del modelo o Determinen el tipo de distribución de probabilidad que ha generado los datos.

Aunque la metodología para realizar el contraste de hipótesis es análoga en ambos casos, distinguir ambos tipos de hipótesis es importante puesto que muchos problemas de contraste de hipótesis respecto a un parámetro son, en realidad, problemas de estimación, que tienen una respuesta complementaria dando un intervalo de confianza (o conjunto de intervalos de confianza) para dicho parámetro. Sin embargo, las hipótesis respecto a la forma de la distribución se suelen utilizar para validar un modelo estadístico para un fenómeno aleatorio que se está estudiando.

Se denomina hipótesis nula  $H_0$  a la hipótesis que se desea contrastar. El nombre de hipótesis nula indica que  $H_0$  representa la hipótesis que mantendremos a no ser que los datos indiquen su falsedad. La hipótesis  $H_0$  nunca se considera probada, aunque puede ser rechazada por los datos.

A partir de una muestra de la población en estudio, se extrae un estadístico (esto es, un valor que es función de la muestra) cuya distribución de probabilidad esté relacionada con la hipótesis en estudio y sea conocida. Se toma entonces el conjunto de valores que es más improbable bajo la hipótesis como región de rechazo, esto es, el conjunto de valores para el que consideraremos que, si el valor del estadístico obtenido entra dentro de él, rechazaremos la hipótesis. La probabilidad de que se obtenga un valor del estadístico que entre en la región de rechazo aún siendo cierta la hipótesis puede calcularse. De esta manera, se puede escoger dicha región de tal forma que la probabilidad de cometer este error sea suficientemente pequeña.

El enfoque actual considera siempre una hipótesis alternativa a la hipótesis nula. De manera explícita o implícita, la hipótesis nula, a la que se denota habitualmente por  $H_0$ , se enfrenta a otra hipótesis que denominaremos hipótesis alternativa y que se denota  $H_1$ . En los casos en los que no se especifica  $H_1$  de manera explícita, podemos considerar que ha quedado definida implícitamente como  $H_0$  es falsa.

Un test de hipótesis se entiende, en el enfoque moderno, como una función de la muestra, corrientemente basada en un estadístico. Supongamos que se tiene una muestra  $X = (X_1, X_2, \dots, X_n)^t$  de una población en estudio y que se han formulado hipótesis sobre un parámetro relacionado con la distribución estadística de la población. Supongamos que se dispone de un estadístico  $T(X)$  cuya distribución con respecto a  $\theta$ ,  $F_\theta(t)$  se conoce. Supongamos, también, que las hipótesis nula y alternativa tienen la siguiente formulación:

$$\begin{aligned} H_0 &: \theta \in \Theta_0 \\ H_1 &: \theta \in \Theta_1 \end{aligned}$$



Un contraste, prueba o test para dichas hipótesis sería una función de la muestra de la siguiente forma:

$$\phi(X) = \begin{cases} 1 & \text{si } T(X) \in \Omega \\ 0 & \text{si } T(X) \notin \Omega \end{cases}$$

Donde  $\phi(X) = 1$  significa que debemos rechazar la hipótesis nula,  $H_0$  (aceptar  $H_1$ ) y  $\phi(X) = 0$ , que debemos aceptar  $H_0$  (o que no hay evidencia estadística contra  $H_0$ ). A  $\Omega$  se la denomina región de rechazo. En esencia, para construir el test deseado, basta con escoger el estadístico del contraste  $T(X)$  y la región de rechazo  $\Omega$ . Se escoge  $\hat{\alpha}$  de tal manera que la probabilidad de que  $T(X)$  caiga en su interior sea baja cuando se da  $H_0$ .

Una vez realizado el contraste de hipótesis, se habrá optado por una de las dos hipótesis,  $H_0$  o  $H_1$ , y la decisión escogida coincidirá o no con la que en realidad es cierta. Se pueden dar los cuatro casos que se exponen en el siguiente cuadro:

|                  | $H_0$ es cierta | $H_1$ es cierta  |
|------------------|-----------------|------------------|
| Se escogió $H_0$ | No hay error    | Error de tipo II |
| Se escogió $H_1$ | Error de tipo I | No hay error     |

Si la probabilidad de cometer un error de tipo I está unívocamente determinada, su valor se suele denotar por la letra griega  $\alpha$ , y en las mismas condiciones, se denota por  $\beta$  la probabilidad de cometer el error de tipo II, esto es:

$$P(\text{escoger } H_1 | H_0 \text{ es cierta}) = \alpha$$

$$P(\text{escoger } H_0 | H_1 \text{ es cierta}) = \beta$$

En este caso, se denomina *Potencia del contraste* al valor  $1 - \beta$ , es decir, a la probabilidad de escoger  $H_1$  cuando esta es cierta

$$P(\text{escoger } H_1 | H_0 \text{ es cierta}) = 1 - \beta$$

Cuando es necesario diseñar un contraste de hipótesis, sería deseable hacerlo de tal manera que las probabilidades de ambos tipos de error fueran tan pequeñas como fuera posible. Sin embargo, Con una muestra de tamaño prefijado, disminuir la probabilidad del error de tipo I,  $\alpha$ , conduce a incrementar la probabilidad del error de tipo II,  $\beta$ .

Usualmente, se diseñan los contrastes de tal manera que la probabilidad sea el 5% (0,05), aunque a veces se usan el 10% (0,1) o 1% (0,01) para adoptar condiciones más relajadas o más estrictas. El recurso para aumentar la potencia del contraste, esto es, disminuir la probabilidad de error de tipo II, es aumentar el tamaño muestral, lo que en la práctica conlleva un incremento de los costes del estudio que se quiere realizar.

Existen distintas aplicaciones de los contrastes de hipótesis. Los contrastes de hipótesis, como la inferencia estadística en general, son herramientas de amplio uso en la ciencia en general. En particular, la moderna filosofía de la ciencia desarrolla el concepto de falsabilidad de las teorías científicas basándose en los conceptos de la inferencia estadística en general y de los contrastes de hipótesis. En este contexto, cuando se desea optar entre dos posibles teorías científicas para un mismo fenómeno (dos hipótesis) se debe realizar un contraste estadístico a partir de los datos disponibles sobre el fenómeno que permitan optar por una u otra.

Las técnicas de contraste de hipótesis son también de amplia aplicación en muchos otros casos, como ensayos clínicos de nuevos medicamentos, control de calidad, encuestas, etcétera.

La prueba chi-cuadrado es cualquier prueba en la que el estadístico utilizado sigue una distribución chi-cuadrado si la hipótesis nula es cierta. La **prueba chi-cuadrado de Pearson** es considerada como una prueba no paramétrica que mide la discrepancia entre una distribución observada y otra teórica (bondad de ajuste), indicando en qué medida las diferencias existentes entre ambas, de haberlas, se deben al azar en el contraste de hipótesis. También se utiliza para probar la independencia de dos variables entre sí, mediante la presentación de los datos en tablas de contingencia.

La fórmula que da el estadístico es la siguiente:

$$\chi^2 = \sum_i \frac{(\text{observada}_i - \text{teórica}_i)^2}{\text{teórica}_i}$$

Cuanto mayor sea el valor de  $\chi^2$ , menos verosímil es que la hipótesis sea correcta. De la misma forma, cuanto más se aproxima a cero el valor de chi-cuadrado, más ajustadas están ambas distribuciones.

Los **grados de libertad gl**, estimador del número de categorías independientes en una prueba particular o experimento estadístico, vienen dados por:

$$gl = (r-1)(k-1)$$

Donde  $r$  es el número de filas y  $k$  el de columnas.

Se acepta la hipótesis nula,  $H_0$ , cuando:

$$\chi^2 < \chi^2_{(1-\alpha)(k-1)}$$

En caso contrario se rechaza. Donde  $t$  representa el valor proporcionado por las tablas, según el nivel de significación estadística elegido.

La **prueba chi-cuadrado de contingencia** o **prueba chi-cuadrado de independencia**, sirve para comprobar la independencia de frecuencias entre dos variables aleatorias,  $X$  e  $Y$ .

Las hipótesis contrastadas en la prueba son:

- Hipótesis nula,  $H_0$ :  $X$  e  $Y$  son independientes.
- Hipótesis alternativa,  $H_1$ :  $X$  e  $Y$  no son independientes (No importa cual sea la relación que mantengan ni el grado de esta).

La condición de independencia es:  $X$  e  $Y$  son independientes si y sólo si para cualquier pareja de valores  $x$  e  $y$  la probabilidad de que  $X$  tome el valor  $x$  e  $Y$  el valor  $y$ , simultáneamente, es igual al producto de las probabilidades de que cada una tome el valor correspondiente.

$$X \text{ e } Y \text{ son independientes} \Leftrightarrow \forall x, y \quad f(x, y) = f(x) \cdot f(y)$$

Por tanto, todo lo que necesitamos serán unas estimaciones de las funciones de probabilidad de ambas variables por separado ( $f(x)$  y  $f(y)$ ) y de la función de probabilidad conjunta ( $f(x, y)$ )

Empezaremos la prueba de independencia tomando una muestra de parejas de valores sobre la que contaremos la frecuencia absoluta con la que aparece cada combinación de valores  $(x_i, y_j)$  o de grupos de valores  $(i, j)$  ( $O_{ij}$ ). La tabla siguiente, en la que se recogen estos datos, es en realidad nuestra estimación de la función de probabilidad conjunta multiplicada por el número total de datos ( $T$ ).

| $X \backslash Y$      | $y_1$    | $y_2$    | ... | $y_i$    | ... | $y_j$    | $F_i = \sum_j O_{ij}$ |
|-----------------------|----------|----------|-----|----------|-----|----------|-----------------------|
| $x_1$                 | $O_{11}$ | $O_{12}$ | ... | $O_{1i}$ | ... | $O_{1j}$ | $F_1$                 |
| $x_2$                 | $O_{21}$ | $O_{22}$ | ... | $O_{2i}$ | ... | $O_{2j}$ | $F_2$                 |
| ...                   | ...      | ...      | ... | ...      | ... | ...      | ...                   |
| $x_i$                 | $O_{i1}$ | $O_{i2}$ | ... | $O_{ii}$ | ... | $O_{ij}$ | $F_i$                 |
| ...                   | ...      | ...      | ... | ...      | ... | ...      | ...                   |
| $x_l$                 | $O_{l1}$ | $O_{l2}$ | ... | $O_{li}$ | ... | $O_{lj}$ | $F_l$                 |
| $C_j = \sum_i O_{ij}$ | $C_1$    | $C_2$    | ... | $C_i$    | ... | $C_j$    | $T$                   |

Para obtener las estimaciones de las funciones de probabilidad marginales debemos sumar por filas y por columnas los valores de las frecuencias conjuntas. Las sumas de filas ( $F_i$ ) son, en cada caso, el número de veces que hemos obtenido un valor de  $X$  ( $x_i$ ) en cualquier combinación con distintos valores de  $Y$ , es decir, son nuestra estima de la función de probabilidad de  $X$  multiplicada por el número total de observaciones; análogamente, las sumas de columnas ( $C_j$ ) son nuestra estima de la función de probabilidad de  $Y$  multiplicada por el número total de observaciones.

## 2.4. Técnicas Soft Computing aplicadas a MDD

El Soft Computing es un término aplicado a un campo dentro de la informática que se caracteriza por el uso de soluciones inexactas para tareas computacionalmente difíciles como la solución de los problemas NP-completos, por lo que una solución exacta no puede ser derivada en tiempo polinómico.

Hasta que en 1994 L.A. Zadeh dio la primera definición de "Soft Computing", la referencia a los conceptos que actualmente ésta maneja solía hacerse de forma atómica, es decir, se hablaba de manera aislada de cada uno de ellos con indicación del empleo de metodologías de lógica difusa. Zadeh propuso la definición de Soft Computing, estableciéndola en los siguientes términos:

"Básicamente, Soft Computing no es un cuerpo homogéneo de conceptos y técnicas. Más bien es una mezcla de distintos métodos que de una forma u otra cooperan desde sus fundamentos. En este sentido, el principal objetivo de la Soft Computing es aprovechar la tolerancia que conllevan la imprecisión y la incertidumbre, para conseguir manejabilidad, robustez y soluciones de bajo costo. Los principales ingredientes de la Soft Computing son la Lógica Fuzzy, la Neuro-computación y el Razonamiento Probabilístico, incluyendo este último a los Algoritmos Genéticos, las Redes de Creencia, los Sistemas Caóticos y algunas partes de la Teoría de Aprendizaje. En esa asociación de Lógica Fuzzy, Neuro-Computación y Razonamiento Probabilístico, la Lógica Fuzzy se ocupa principalmente de la imprecisión y el Razonamiento Aproximado; la Neuro-computación del aprendizaje, y el Razonamiento Probabilístico de la incertidumbre y la propagación de las creencias".

Queda claro así que el Soft Computing no está definido precisamente, sino que en una primera aproximación se define por extensión, por medio de distintos conceptos y técnicas que intentan superar las dificultades que surgen en los problemas reales que se dan en un mundo que es impreciso, incierto y difícil de categorizar.

Aunque ha habido varios intentos de ajustar más esta definición, no han sido muy fructíferos. Así, por ejemplo a la vista de la dificultad de dar una nueva definición del campo de una manera exacta y consensuada, y de la mayor sencillez de hacerlo por medio de sus características, los autores proponen la siguiente definición de trabajo, que vuelve a ser de tipo descriptivo: "Cualquier proceso de computación que expresamente incluya imprecisión en los cálculos en uno o más niveles, y que permita cambiar (disminuir) la granularidad del problema o suavizar los objetivos de optimización en cualquier etapa, se define como perteneciente al campo de la Soft Computing".

El Soft Computing se convirtió en una rama formal de la informática a principios de los años 90. Las primeras aproximaciones informáticas sólo podían modelar y analizar con precisión sistemas relativamente simples. Los sistemas más complejos que aparecían en biología, medicina, humanidades y administración de empresas, campos en los que muchos cálculos resultaban inmanejables con los métodos analíticos y matemáticos convencionales. Es necesario decir que la simplicidad y la complejidad de los sistemas son relativos, y muchos modelos matemáticos convencionales han sido muy productivos.

Como técnicas de Soft computing se incluyen las redes neuronales, sistemas difusos, o sistemas basados en la teoría del caos. Dentro de la bioinformática se pueden encontrar varias técnicas Soft-Computing, tales como algoritmos evolutivos, optimización de colonias de hormigas o la inteligencia de enjambre. En el campo de la probabilidad estadística se pueden destacar las redes Bayesianas como ejemplo de Soft-Computing.

Generalmente, las técnicas de soft computing se asemejan más a los procesos biológicos que a las técnicas matemáticas tradicionales, que se basan principalmente en sistemas formales lógicos, tales como Lógica proposicional y lógica de predicados, o basados en análisis numérico asistido por ordenador (como en el método de los elementos finitos). Las técnicas de Soft computing intentan complementarse unas a otras.

A diferencia de las técnicas de hard computing (computación tradicional), que buscan la exactitud y la verdad total, las técnicas de soft computing explotan la tolerancia de la imprecisión, la verdad parcial y la incertidumbre para un problema específico. Otra diferencia que contrasta aparece al considerar que el razonamiento inductivo juega un papel más destacado en soft computing que en hard computing.

La Lógica difusa (fuzzy logic+en inglés) se aplica con éxito para resolver problemas de control. Fue introducida por primera vez por L.A. Zadeh que trabajaba con estos problemas y se dio cuenta que, utilizando ecuaciones diferenciales, en muchas ocasiones no se llegaba a soluciones prácticas, bien por las simplificaciones que había que introducir, bien por los muchos cálculos que era preciso realizar, por lo que precisó utilizar nuevas herramientas.

Un subconjunto clásico se conoce al saber los elementos de un conjunto referencial que pertenecen a él. La idea de Zadeh es proponer un grado de pertenencia. Mientras que en un conjunto clásico un elemento, o pertenece o no pertenece, ahora en la lógica difusa, puede pertenecer con un cierto grado. Esta flexibilidad permite tratar con problemas de incertidumbre y con problemas de conocimiento incompleto. Un ejemplo muy utilizado es el subconjunto de ~~de~~ personas altas. Si una persona mide 2 m es alta, y si mide 1 mm menos que una persona alta, sigue siendo alta. Con este razonamiento se llega a que alguien que mida 1.2 m es una persona alta. Este conjunto, el de las personas altas, no tiene unas fronteras claramente definidas. Y lo mismo ocurre con muchos otros muchos conjuntos, el de jóvenes, rubios, guapos, que saben inglés... con todos aquellos cuya medida requiera una cierta subjetividad.

Un conjunto clásico viene dado por su función característica,  $f$ , que asigna un cero a un elemento que no pertenece al conjunto y un 1 al que si pertenece:  $f: X \rightarrow \{0, 1\}$ . Un conjunto difuso,  $A$ , se define con su función de pertenencia, que a cada elemento le asigna un valor comprendido entre cero y uno:

$$A: X \rightarrow [0, 1].$$

Se puede observar que las funciones de transformación son precisamente funciones de un conjunto referencial, en el que se definen las magnitudes en unidades heterogéneas, y el intervalo cerrado  $[0, 1]$ , es decir, pueden ser consideradas como conjuntos difusos:  $M: X \rightarrow [0, 1]$ . Al transformar la magnitud de cada impacto de unidades heterogéneas a unidades homogéneas se puede considerar que se utiliza un conjunto difuso, lo que permite usar todos los instrumentos de esta teoría.

La lógica clásica, la teoría de conjuntos clásica o la teoría de probabilidad pueden no ser adecuadas para tratar la imprecisión, la incertidumbre, la no especificidad, la vaguedad, la inconsistencia y la complejidad del mundo real. Esto motiva el nacimiento de los conjuntos difusos y las lógicas borrosas y explica su papel en la reestructuración de los fundamentos de las teorías científicas y sus aplicaciones, por lo que se están produciendo grandes avances tanto en áreas teóricas como en gran variedad de aplicaciones.

La teoría de la probabilidad sólo es capaz de representar uno de los tipos de incertidumbre que se basa en la aleatoriedad, no en la imprecisión de la información. L.A. Zadeh en 1965 escribe su artículo en el que introduce una teoría sobre unos objetos, los conjuntos difusos, que son conjuntos de frontera no precisa y cuya función de pertenencia indica un grado. En la esfera de los predicados subjetivos, y por tanto imprecisos, la teoría de conjuntos clásica se enfrenta con obstáculos difíciles de superar.

Las lógicas borrosas necesitan generalizar las conectivas entre conjuntos borrosos. Los conectivos lógicos AND, OR y NOT, y las operaciones entre conjuntos intersección, unión y negación se generalizan respectivamente mediante normas triangulares, conormas triangulares y negaciones.

Cuando se pretende utilizar las relaciones borrosas para efectuar inferencias de razonamiento aproximado de la forma:

Si  $x$  es  $P$  entonces  $y$  es  $Q$

$x$  es  $P_q$

---

$y$  es  $Q_q$

Parece necesario estudiar las propiedades algebraicas de las relaciones borrosas. Las más utilizadas son la reflexividad, simetría y T-transitividad. Una relación borrosa con estas tres propiedades es una T-indistinguibilidad, que generaliza a una relación de equivalencia. Sus aplicaciones son diversas, como la comparación y la clasificación, y es utilizada también en el aprendizaje inductivo automático. Es interesante estudiar las indistinguibilidades como complemento o negación de distancias, aprovechando los conocimientos sobre espacios métricos y espacios métricos generalizados.

Dentro de la rama de lógica difusa, uno de los procedimientos de decisión que se utilizan son los métodos de decisión multicriterio. Estos métodos de ayuda a la decisión dan a quien debe tomarlas algunas herramientas para que sea capaz de avanzar en la resolución de problemas de decisión donde muchos puntos de vista deben ser tenidos en cuenta (frecuentemente contradictorios).

No existe, en general, ninguna decisión (solución, acción) la cual sea mejor simultáneamente desde todos los puntos de vista. La optimización no tiene mucho sentido, en contraste a las técnicas clásicas de investigación de operaciones.

En los métodos de decisión multicriterio se modeliza el problema de decisión mediante la definición previa de una serie de alternativas (candidatos) entre las que se desea escoger la más adecuada por lo que son excluyentes. Para cada alternativa se evalúa un conjunto de criterios ponderados según la importancia que tengan. A partir de la ponderación y la evaluación de los criterios se aplica un procedimiento que devuelve un valor para cada alternativa, por lo que la aplicación devuelve realmente una ordenación de las alternativas, de modo que la mejor valorada corresponde con la solución del problema.

Para mostrar los pasos del método utilizado y los resultados obtenidos se plantea la siguiente pregunta; De los 3 grupos de edad, ¿Cuál de ellos se espera que aumente el consumo de marcas de distribuidor (MDD)?

El problema de decisión multicriterio en este caso posee  $m=3$  alternativas (menor de 25 años, entre 25 y 50 años, mayor de 50 años) y  $n=4$  criterios correspondientes a las variables sobre la compra de productos de MDD, compra con conocimiento del consumidor de la MDD, aumento del consumo y mantenimiento del consumo.

En la tabla 4 se puede ver los resultados obtenidos directamente de la BD de consumidores:

| OLA 14  | Sin Conocimiento | Con conocimiento | Más Consumo | Igual Consumo |
|---------|------------------|------------------|-------------|---------------|
|         | Peso = 0,3       | Peso = 0,3       | Peso = 0,2  | Peso = 0,2    |
| < 25    | 0,86             | 0,42             | 3           | 3             |
| 25 - 50 | 0,76             | 0,59             | 8           | 42            |
| > 50    | 0,80             | 0,61             | 4           | 21            |

Tabla4. Resultados obtenidos directamente de la B.D

Esta misma tabla es la matriz de decisión y se muestra también el peso que se le ha asignado a cada uno de los criterios atendiendo la relevancia para el estudio. Para aplicar el procedimiento MCDM y previo a la normalización de los resultados de la matriz de decisión, se asocia a cada elemento de dicha matriz su número triangular correspondiente.

Con los números triangulares, conseguimos representar el dato como una zona de incertidumbre alrededor del valor tomado. Por simplicidad en este caso se ha tomado un intervalo fijo alrededor del punto con una variación de un 10%.

La Tabla 5 muestra la matriz de decisión con las tuplas que representan a los números triangulares donde además se ha realizado una normalización en una escala 0-1. Por último, la Tabla 6 es el resultado de aplicar los pesos a los datos de la matriz de decisión.

| OLA 14        | Sin Conocimiento<br>Peso = 0,3 | Con conocimiento<br>Peso = 0,3 | Más Consumo<br>Peso = 0,2 | Igual Consumo<br>Peso = 0,2 |
|---------------|--------------------------------|--------------------------------|---------------------------|-----------------------------|
| Menor de 25   | (0,77;0,86;0,95)               | (0,38;0,42;0,46)               | (0,27;0,30;0,33)          | (0,27;0,30;0,33)            |
| Entre 25 y 50 | (0,68;0,76;0,84)               | (0,53;0,59;0,65)               | (0,12;0,13;0,14)          | (0,62;0,69;0,76)            |
| Mayores de 50 | (0,72;0,80;0,88)               | (0,55;0,61;0,67)               | (0,10;0,11;0,12)          | (0,54;0,60;0,66)            |

Tabla5. Matriz de decisión

| ESTUDIO DE LA OLA14 | Compra<br>Peso=0,3 | Compra sobre conocimiento<br>Peso=0,3 | Consumen más que el año anterior<br>Peso=0,2 | Consumen igual que el año anterior<br>Peso=0,2 |
|---------------------|--------------------|---------------------------------------|--|--|
| Menor de 25         | (0,23;0,26;0,29)   | (0,11;0,13;0,14)                      | (0,05;0,06;0,07)                             | (0,05;0,06;0,07)                               |
| Entre 25 y 50       | (0,20;0,23;0,25)   | (0,16;0,18;0,20)                      | (0,02;0,03;0,03)                             | (0,12;0,14;0,15)                               |
| Mayores de 50       | (0,22;0,24;0,26)   | (0,17;0,18;0,20)                      | (0,02;0,02;0,02)                             | (0,11;0,12;0,13)                               |

Tabla6. Matriz de decisión aplicando los pesos

En la última fase del método del vértice se debe calcular la distancia de cada uno de los números triangulares a los puntos ideales positivo y negativo respectivamente. La alternativa preferida será la que tenga la menor distancia al ideal positivo (1,1,1) y la mayor al ideal negativo (0,0,0).

Las tablas 7 y 8 muestran los resultados de calcular las distancias a los ideales positivo y negativo respectivamente.

| OLA 14  | Sin Conocimiento<br>Peso = 0,3 | Con conocimiento<br>Peso = 0,3 | Más Consumo<br>Peso = 0,2 | Igual Consumo<br>Peso = 0,2 |
|---------|--------------------------------|--------------------------------|---------------------------|-----------------------------|
| < 25    | 0,74                           | 0,87                           | 0,94                      | 0,94                        |
| 25 - 50 | 0,77                           | 0,82                           | 0,97                      | 0,86                        |
| > 50    | 0,76                           | 0,82                           | 0,98                      | 0,88                        |

Tabla7. Resultados de calcular las distancias a los ideales positivos

| OLA 14  | Sin Conocimiento<br>Peso = 0,3 | Con conocimiento<br>Peso = 0,3 | Más Consumo<br>Peso = 0,2 | Igual Consumo<br>Peso = 0,2 |
|---------|--------------------------------|--------------------------------|---------------------------|-----------------------------|
| < 25    | 0,26                           | 0,13                           | 0,06                      | 0,06                        |
| 25 - 50 | 0,23                           | 0,18                           | 0,03                      | 0,14                        |
| > 50    | 0,24                           | 0,18                           | 0,02                      | 0,12                        |

Tabla8. Resultados de calcular las distancias a los ideales negativos



El criterio de selección de candidato se realiza aplicando una fórmula que da como resultado una ordenación total de los datos. Este criterio utiliza la idea de agregación de las distancias a los ideales calculadas anteriormente:

$$p_i = \frac{d_i^- + n - d_i^+}{2n}$$

Los elementos  $d_i^-$  y  $d_i^+$  se obtendrán de la suma de los elementos  $d_i$  correspondientes a cada campo, es decir, para cada fila, su componente correspondiente será la resultante de la suma de sus distancias a los ideales, para cada criterio. Así, por ejemplo, para OLA 14 en el campo < 25 su  $d_i^-$  será la suma  $0,74+0,87+0,94+0,94$ .  $d_i = 3,49$ .

La Tabla 9 muestra los resultados obtenidos. Estos resultados son ya las valoraciones de las tres alternativas en estudio.

| OLA 14 | Menor de 25 | Entre 25 y 50 | Mayor de 50 |
|--------|-------------|---------------|-------------|
| $P_i$  | 0,13        | 0,15          | 0,14        |

Tabla9. Resultado final

La interpretación de los resultados es la siguiente: Se espera que el sector poblacional que mayor consumo realizará de MDD sea entre 25 y 50 años, y el que menor consumo hará será el de menores de 25. Por lo tanto, las compañías deben reforzar mediante campañas publicitarias u otras técnicas de mercados el consumo de éste último sector o bien aprovechar la buena acogida del sector entre 25 y 50 años para crear productos específicos que reviertan en un beneficio para las empresas.

El resultado anterior se ha basado únicamente en los datos procedentes de la OLA 14 a modo de ejemplo para mostrar la teoría que hay detrás de la ejecución del programa.



## **3. Desarrollo del proyecto**

### **3.1. Introducción**

La globalización de mercados en el entorno en el que vivimos ha provocado nuevas formas de comercialización de productos más competitivas y dirigidas a sectores poblacionales más específicos a la vez que se han aparecido nuevos formatos de producción, como es el caso de las Marcas de Distribuidor (MDD).

El fenómeno de las MDD también conocidas como Store Brands (SB) se ha desarrollado a lo largo de los últimos 30 años internacionalmente aunque fundamentalmente en sociedades occidentales. Este desarrollo se ha ido adaptando a la situación económica de cada país y es muy interesante el estudio de la evolución de las MDD en los distintos mercados a lo largo de los últimos años, para ello es interesante estudiar los datos anuales almacenados, conocidos como OLAS.

La Universidad Complutense de Madrid y en concreto el Departamento de Comercialización e Investigación de Mercados ha desarrollado durante más de 20 años una base de datos sobre resultados obtenidos por encuesta a pie de calle, datos relativos al consumo de Marcas de Distribuidor de las enseñas más importantes con comercialización en centros de distribución de la Comunidad de Madrid.

Como consecuencia, se ha obtenido una gran base de datos con más de 4000 encuestas realizadas a consumidor y más de 56000 registros de productos. Aunque los datos están originalmente almacenados en ficheros Excel bien estructurados el primer problema a abordar es la integración de los datos en una única base de datos analizable desde el punto de vista estadístico. Para ello se ha de tener en cuenta los cambios que a lo largo de los años han sufrido los formularios de encuesta al consumidor y los diferentes formatos de presentación del producto que han ido surgiendo o desapareciendo, entre otras consideraciones.

El primer paso por lo tanto ha sido la elaboración de un repositorio para los datos sobre compradores y otro para los datos sobre productos que respectivamente proceden de los formularios sobre compradores y sobre productos. El primer formulario se ha ido obteniendo mediante encuesta a pie de calle a personas que acaban de realizar una compra en una de las enseñanzas participantes en el estudio y el segundo formulario se ha obtenido mediante anotaciones que los colaboradores (estudiantes de la facultad de CC. Económicas y Empresariales) han recogido en las diferentes enseñanzas a lo largo de los años como trabajo de campo.

Debido al interés de obtener análisis estadísticos sobre los datos y tras estudiar las posibilidades de la herramienta SPSS para el almacenamiento, se decidió la utilización de SPSS también como base de datos. Sin embargo, el interés por responder a determinadas preguntas sobre mercadotecnia y comportamiento del consumidor, ha derivado en la realización de una herramienta software que sirva como interfaz entre el experto en mercados pero desconocedor de herramientas de análisis estadístico. Uno de los objetivos fundamentales de esta herramienta es servir como culminación del trabajo de campo realizado por los estudiantes, pudiendo analizar y obtener resultados sobre los datos que previamente han recogido.

Se pretende realizar una aplicación en C# de fácil utilización y con una interfaz amigable, para la realización de análisis de encuestas de marcas de distribuidor (SBDataAnalysis).

El sistema tomará como datos de entrada, unas bases de datos con la información de las diferentes encuestas de marcas de distribuidor.

El sistema tendrá como función principal la elaboración de diversos estudios con respecto a las respuestas de las encuestas. Existen diversos tipos de estudios como pueden ser estudios de evolución de respuestas, estudios de comparación de respuestas por enseñanza o formato y estudios de independencia entre diversas respuestas. De estos estudios se crean unos informes en formato Excel en el que se detalla los resultados obtenidos. Además de los informes, si procede, se mostrarán unos gráficos con los resultados obtenidos.

Además de estos estudios, el sistema realiza un análisis de comparación de consumo con respecto a los rangos de edad de los encuestados y otra, utilizando técnicas de Soft-Computing y en particular técnicas de decisión multicriterio.

También contará con la función de mostrar la información almacenada en las diferentes bases de datos de las encuestas y con un sistema de ayuda, en el que se le informa al usuario del modo de uso de las diferentes opciones del sistema y del significado de las diversas variables de las bases de datos.

Este proyecto está pensado para el departamento MDD en su labor de estudiar el comportamiento y la evolución de los consumidores con respecto a las Marcas de distribuidor.

Les ayudará a crear informes sobre diferentes estudios de la evolución de diversos datos de las encuestas con respecto a diferentes variables (olas, enseñanzas, rango de edad, etc.). En estos informes los miembros del departamento MDD podrán ver y analizar diferentes variables estadísticas sobre estos estudios, que les ayudará a comprender mejor el comportamiento de los consumidores de MDD. Además de crear los informes, se mostrarán unos gráficos de los estudios que les hará más fácil ver la evolución o los valores de las variables que intervienen en el estudio.

Además de los informes de evolución, se crean informes de estudios de dependencia entre dos variables, que les informará si dos variables tienen dependencia entre sí y, por tanto, si los valores de cada variable tienen un patrón en común. Estos estudios aclararán si el cambio en los valores de una variable afecta directamente sobre los valores obtenidos en la otra variable.

El proyecto comprende una parte de comparación de Olas, en la que se hace una comparación entre los diferentes rangos de edad que recoge la encuesta (menos de 25 años, entre 25 años y 50 años y mayor de 50 años) sobre su consumo de MDD y nos indica el orden de consumo de estos rangos de edad. Con esta información los miembros del departamento MDD pueden ver cuál es el rango de edad más proclive a consumir MDD en las diferentes Olas de encuestas y, por tanto, ver su evolución. Sabiendo cuál es el rango que más tendencia tendrá a consumir, saben cuál es el rango de edad al que deberían ir dirigidas los diferentes mecanismos de publicidad de la que dispone los fabricantes de MDD y así aumentar su consumo.

Este proyecto también servirá para los alumnos de las diferentes carreras que ayudan al departamento MDD recogiendo encuestas, ayudándoles a introducir la información a un soporte electrónico etc. Podrán ver los informes creados, tanto de los estudios de la evolución como los estudios de dependencia, además de la comparación de Olas; sacando conclusiones de esos estudios y entendiendo mejor el mundo de las MDD y los factores que afectan de forma considerable a este tipo de productos.

Las diversas conclusiones que se pueden sacar de los diferentes estudios del proyecto podrían ayudar a un fabricante de MDD a comprender mejor la evolución de los gustos de los consumidores y actuar en consecuencia cambiando diferentes aspectos del producto, como puede ser el envase, la publicidad, su situación en el establecimiento etc.; aumentando su consumo y beneficio. En contraposición, las conclusiones sacadas en los diferentes estudios del proyecto también podrían ayudar a un fabricante de Marcas de Fabricante a contraatacar a las MDD, sabiendo cuál es la evolución de las MDD en los gustos de los consumidores y actuar en consecuencia; haciendo que no aumente, o aumente lo mínimo posible, el consumo de las MDD.

Los principales usuarios del proyecto son los miembros del departamento MDD, ya que se ajusta a las necesidades que ellos tienen con respecto al estudio de sus encuestas. El proyecto se ha construido a partir de entrevistas que se han hecho con ellos en las cuales informaron sobre qué estudios les gustaría realizar y de ahí se sacaron las conclusiones que ellos necesitan para la realización de estudios más grandes y artículos que las personas de dicho departamento hacen con las conclusiones de estos estudios.

Como posibles usuarios también están los alumnos que ayudan al departamento MDD en la realización de la recogida de información de los productos y las encuestas en la puerta de los centros a los consumidores. Los alumnos podrían ayudar al departamento MDD a sacar conclusiones de los resultados de los diferentes estudios que se realizan en el proyecto.

En definitiva los posibles usuarios de este proyecto son todas aquellas personas que tengan un especial interés en saber cómo evolucionan los gustos y la puntuación de las características de las MDD en los consumidores. Estos usuarios estarían tanto dentro del entorno académico, profesores y alumnos; como dentro del sector privado, fabricantes de MDD, fabricantes de MF, agencias de publicidad, distribuidoras etc.; todos estos relacionados con las MDD que buscarían una manera de aumentar o disminuir, en el caso de la competencia, el consumo de estos productos.

El conocimiento para el proyecto respecto al tema de marcas de distribuidor era muy limitado. Hasta la primera reunión no se supo que el departamento de Investigación y Comercialización de Mercados de la Universidad Complutense de Madrid llevaba realizando durante muchos años encuestas sobre las marcas de distribuidor.

Al principio hubo que leer bastante información sobre marcas de distribuidor y leer unas pequeñas guías que el departamento MDD daba a los usuarios que participaban en la elaboración de las encuestas para poder entender qué preguntaban a la puerta de los supermercados o qué información apuntaban con respecto a los productos de marcas de distribuidor expuestas en las estanterías de los supermercados.

Después de obtener información teórica con respecto a las marcas de distribuidor y a las encuestas que se realizan, se facilitaron los archivos Excel donde se almacenaba la información. Hubo que analizar su contenido y ver como estaba organizada la información.

Una vez analizados los datos de las encuestas, se tuvo una reunión con el departamento MDD donde informaron más o menos de las funciones que les gustaría que el programa hiciera. Escribieron un documento en el que se explicaban los principales estudios que ellos realizan, el cuál se analizó y se buscó la manera de poder hacer dichos estudios en los programas.

En el caso de la parte de comparación de consumo por rango de edad y ola, nuestra tutora, Maria Victoria López López, ayudo a la hora de orientar el estudio y de aplicar las técnicas de Soft-Computing. Además de orientar, facilitó bastante información de la materia que resulto vital para poder realizar el estudio.

Además de obtener formación en el campo de las marcas blancas, también fue necesario como parte de la inmersión dentro del mundo de los conceptos y estudios estadísticos, asistir a un curso organizado por la Universidad Complutense en el cual se enseñó a trabajar con SPSS a nivel de usuario. Este cursillo fue fundamental a la hora de la realización del proyecto ya que con él se adquirieron gran parte de los conocimientos que luego se han ido empleando en la elaboración de este proyecto, así como la familiarización con conceptos que se han tenido que usar y que han hecho que se avanzara con mayor rapidez en ciertos momentos.

## 3.2. Tecnologías Utilizadas

Antes de iniciar este proyecto no había una premisa concreta sobre cómo o con qué herramientas iba a llevarse a cabo. Las únicas premisas eran que debía tener una sólida componente estadística y ser de sencillo manejo.

Así pues en un primer momento lo lógico era conocer las herramientas software estadísticas que se podían tener al alcance, las recomendaciones de la directora del proyecto fueron encaminadas hacia tres:

- SPSS
- SAS
- Clementine.

De estos dos primeros, se realizaron cursos de la Universidad Complutense para conocer su funcionamiento. Si bien estos cumplieron su función, que era enseñar al alumno a usar la herramienta, no cumplieron plenamente las expectativas que se tenían, ya que no ofrecían modo alguno de cómo aplicarlo al proyecto que se había escogido. Era el momento de ponerse a investigar.

La fuente principal a la que se recurrió fue Internet. Estaba claro que era fundamental poder aprovechar los recursos que pudiera ofrecer un programa estadístico para que los esfuerzos de implementación del proyecto fueran encaminados a personalizar el trabajo para las necesidades del grupo de investigación de MDD y no a implementar toda una serie de tediosas funciones estadísticas que harían que el proyecto se desviara del objetivo principal de diseñar un software muy específico.

Entonces fue cuando se encontró la funcionalidad que se se buscaba, SPSS, una de las herramientas que se tenía a disposición. Tenía unos módulos complementarios de programabilidad, esto es, permite usando ciertos lenguajes de programación añadir mejoras o construir software basado en él.

**SPSS**, anteriormente conocido como **%Statistical Package for the Social Sciences+**, es una compañía estadounidense fundada en 1968 por Norman H Nie, politólogo y profesor emérito de la Universidad de Chicago.

Su producto estrella también se llama SPSS, aunque durante mucho tiempo también fue conocido como PASW (Predictive Analytics Software).

SPSS es una potente herramienta estadística, que permite hacer una ingente cantidad de estudios y que ha sido empleado desde su aparición por gobiernos, compañías de seguros, compañías médicas etcétera... para sus investigaciones.

Para comprender mejor lo que implica SPSS en el mercado de software estadístico, basta con ver las cifras con las que la compañía IBM adquirió SPSS en julio de 2009. El gigante informático desembolsó 1200 millones de dólares, esto es, unos 960 millones de euros.

Aparte de los motivos anteriormente expuestos, SPSS cumple otro requisito que hizo que el proyecto se decantara por la idea de basarse en él para su realización, y es que es un software con el que también cuenta el grupo de trabajo de MDD. Así pues, no era necesario que se tuvieran que realizar laboriosos cambios ni migraciones masivas.

Pero la característica que hizo que se decantara por el uso de SPSS fue por encima de todas su módulo de programabilidad. Este complemento permite a cualquier desarrollador crear sus propios complementos para SPSS o bien permite construir herramientas software basadas en SPSS, como es nuestro caso.

Dentro del módulo de programabilidad existen dos opciones de trabajo, usar un editor del propio SPSS para programar, o bien descargar unas bibliotecas compatibles bien con Python, bien con .Net o bien con R. Como se tenía que desarrollar un software cuya interfaz gráfica debía contar con un peso notable, se decidió que era mejor idea incluir estas bibliotecas en el editor correspondiente y trabajar desde él con mayor libertad.

SPSS introdujo en su versión 14 una interesantísima novedad, poder mejorar el software mediante la programación del usuario, era el año 2005 y sólo se podía realizar con Python, aunque posteriormente se fue incluyendo soporte para .Net y en las últimas versiones también se incluye al entorno libre de ~~R~~ Rq

En momentos iniciales del proyecto, y una vez decidido que el software ideado se basaría en SPSS, la prioridad era aprender a utilizar el módulo de programabilidad, pero ésta, lejos de ser algo popular entre los usuarios de SPSS era una característica muy poco usada y por lo tanto la cantidad de información disponible era a su vez bastante escasa.

La única fuente realmente útil era la propia página web de la compañía, que tenía una serie de presentaciones realizadas en power point explicando de manera escueta algunos ejemplos de uso del módulo. Además del problema de la falta de documentación para trabajar había que añadir que Python era un lenguaje totalmente nuevo y había que empezar desde cero.

Afortunadamente, en el mes de enero la compañía dio un gran empujón al plug-in de programabilidad liberando una gran cantidad de documentación y ejemplos para .Net, prácticamente todo en visual basic. En este momento y paralelamente la directora de proyecto se puso en contacto con un profesor de la universidad experto conocedor de SPSS, Juan Carlos Martínez, que aconsejó el uso de las tecnologías .Net para el desarrollo del proyecto.

.Net es un framework desarrollado por Microsoft muy orientado a la web, pero que también permite realizar trabajos de escritorio. En él están comprendidos varios lenguajes con Visual Basic, C#, Delphi, C++, o J#.



Conociendo el rango de posibles lenguajes con el que realizar la implementación se optó por C# debido fundamentalmente a su potencia, pero también a que era un lenguaje relativamente desconocido pero con el que rápidamente se podría adquirir destreza debido a los conocimientos de otros lenguajes como Java o C++, muy parecidos tanto en sintaxis como en estructura a C #.

El plug-in de SPSS es de libre distribución, y se puede descargar desde la propia página web de la compañía <http://www.spss.com/devcentral/index.cfm?pg=plugins> según la versión del programa con la que se quiera trabajar y el lenguaje que se requiera de los disponibles. Es condición indispensable tener previamente instalado el software de SPSS para que este módulo tenga alguna utilidad.

Este plug-in integra en el SPSS una serie de bibliotecas que permiten conectar un programa creado bajo cualquiera de los lenguajes disponibles, Python, C# o R con SPSS y hacer que éste se ejecute y proporcione resultados. En otras palabras, permite mandar órdenes de ejecución a SPSS de la misma forma que si se estuviera usando el software propiamente dicho, pero desde una herramienta independiente.

Esto se consigue ejecutando el programa SPSS en backend, esto es, el proceso asignado a SPSS se inicia normalmente, pero el usuario final no se percata de ello, ya que no se muestra notificación alguna de que ha sido iniciado, y es a partir de este momento en el que comienzan a interactuar entre SPSS y el software que se quiera desarrollar.

Para poder hacer funcionar bien el módulo de programabilidad, ha habido que crearse un objeto de la clase Processor a la que se pasa la ruta donde se encuentra instalado el SPSS. Processor es una clase incluida en las bibliotecas del módulo de programabilidad y que lleva el peso de interconexión de nuestra herramienta con SPSS.

Una vez creado este objeto, es el momento de inicializar las comunicaciones, para ello la clase Processor tiene un método StartSPSS que abre el flujo. Análogamente también existe el método CloseSPSS, que cierra dicho flujo. Cualquier comunicación que se quiera hacer con SPSS debe estar incluida entre estas dos instrucciones, de otra manera, no tendrán efecto alguno.

El funcionamiento de este módulo es relativamente sencillo. Cuando se usa el software de SPSS de forma normal, esto es, usar el propio software en sí para realizar pruebas, estudios u operaciones, para cada actividad que se realiza se abre una pantalla de resultados llamada Visor, en la que se puede consultar el contenido de un fichero log que almacena las operaciones que se han ido realizando y los resultados obtenidos, como se puede comprobar en la figura 8. Estas operaciones tienen un formato como este:

```
FREQUENCIES VARIABLES=IDCENTRO P2A  
  /STATISTICS=MEAN  
  /ORDER=ANALYSIS.
```

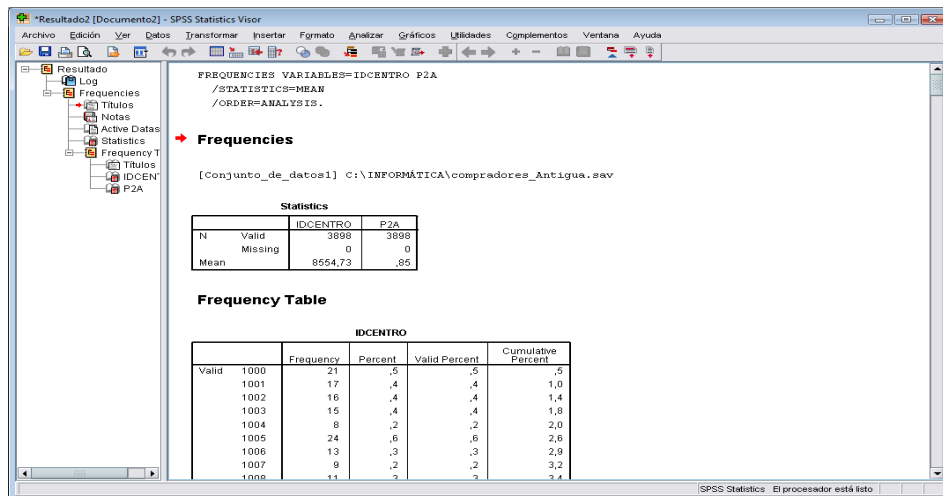


Figura 8. Tabla de frecuencia

Este texto representa que se quiere hacer un análisis de 2 variables, idcentro y p2A con la media de sus valores. Pues si se quisiera realizar lo mismo, pero utilizando la herramienta y ejecutando SPSS en backend, lo único que habría que hacer es construir un String que contenga la cadena anteriormente mostrada, y mandarla a SPSS mediante el método Submit(cadena).

```
String ord = "FREQUENCIES VARIABLES=IDCENTRO P2A";
ord += "/STATISTIC=MEAN";
ord += "/ORDER=ANALYSIS";
Processor.Submit(ord);
```

Esta orden es enviada a SPSS que la ejecuta, en este caso realiza la tabla de frecuencia entre IdCentro y P2A, aunque el usuario no se percataría y tampoco aparecería en la ventana de Resultados de SPSS en el caso de tenerla abierta. Una forma de recoger los resultados sería exportarlos a un archivo para que el usuario pueda observarlos, en este caso, además del estudio debemos incorporar una orden de exportación, en nuestro caso a un archivo de tipo Excel.

```
String begin = "OMS SELECT TABLES /DESTINATION FORMAT=XLS";
begin += "OUTFILE='C:/estudios/resultado.xls'.";

String analysis = "FREQUENCIES VARIABLES=IDCENTRO P2A";
analysis += "/STATISTIC=MEAN";
analysis += "/ORDER=ANALYSIS";

string end = "OMSEND.";

string[] ord = new string[3];
prob[0] = begin;
prob[1] = analysis;
prob[2] = end;

SPSS_Processor.Submit(ord);
```

Una vez mandada la orden se exportaría el análisis a un documento llamado *resultado.xls* y se almacenaría en la carpeta *C:\estudios\*.

En la figura 9 se puede comprobar como es el fichero que se genera.

Microsoft Excel - resultados

Archivo Edición Ver Insertar Formato Herramientas Datos Equipq Ventana ?

Escriba una pregunta

Nueva lista Obtener elementos de trabajo Publicar Actualizar Configurar lista Elegir columnas Vinculos y datos adjuntos

| A1 | A                      | B                                      | C   | D       | E             | F                  | G | H | I | J | K | L |
|----|------------------------|--|---|---------|---------------|--------------------|---|---|---|---|---|---|
| 16 | Missing Value Handling | Definition of Missing                  | User-defined missing values are treated as missing.                           |         |               |                    |   |   |   |   |   |   |
| 17 |                        | Cases Used                             | Statistics are based on all cases with valid data.                            |         |               |                    |   |   |   |   |   |   |
| 18 | Syntax                 |  | FRECUENCIAS<br>VARIABLES=IDCENTRO P2A<br>/STATISTICS=MEAN<br>/ORDER=ANALYSIS. |         |               |                    |   |   |   |   |   |   |
| 19 | Resources              | Processor Time                         | 0:00:00.062   |         |               |                    |   |   |   |   |   |   |
| 20 |                        | Elapsed Time                           | 0:00:00.152   |         |               |                    |   |   |   |   |   |   |
| 21 |                        |  |   |         |               |                    |   |   |   |   |   |   |
| 22 |                        |  |   |         |               |                    |   |   |   |   |   |   |
| 23 | [Conjunto_de_datos1]   | C:\INFORMÁTICA\compradores_Antigua.sav |   |         |               |                    |   |   |   |   |   |   |
| 24 |                        |  |   |         |               |                    |   |   |   |   |   |   |
| 25 | <b>Statistics</b>      |  |   |         |               |                    |   |   |   |   |   |   |
| 26 |                        |  | IDCENTRO  | P2A     |               |                    |   |   |   |   |   |   |
| 27 | N                      | Valid                                  | 3898  | 3898    |               |                    |   |   |   |   |   |   |
| 28 |                        | Missing                                | 0   | 0       |               |                    |   |   |   |   |   |   |
| 29 | Mean                   |  | 8554.73   | .85     |               |                    |   |   |   |   |   |   |
| 30 |                        |  |   |         |               |                    |   |   |   |   |   |   |
| 31 |                        |  |   |         |               |                    |   |   |   |   |   |   |
| 32 | <b>Frequency Table</b> |  |   |         |               |                    |   |   |   |   |   |   |
| 33 |                        |  |   |         |               |                    |   |   |   |   |   |   |
| 34 |                        |  | IDCENTRO  |         |               |                    |   |   |   |   |   |   |
| 35 |                        |  | Frequency   | Percent | Valid Percent | Cumulative Percent |   |   |   |   |   |   |
| 36 | Valid                  | 1000                                   | 21  | .5      | .5            | .5                 |   |   |   |   |   |   |
| 37 |                        | 1001                                   | 17  | .4      | .4            | 1.0                |   |   |   |   |   |   |
| 38 |                        | 1002                                   | 16  | .4      | .4            | 1.4                |   |   |   |   |   |   |
| 39 |                        | 1003                                   | 15  | .4      | .4            | 1.8                |   |   |   |   |   |   |
| 40 |                        | 1004                                   | 8   | .2      | .2            | 2.0                |   |   |   |   |   |   |
| 41 |                        | 1005                                   | 24  | .6      | .6            | 2.6                |   |   |   |   |   |   |

Sheet1

Figura 9. Archivo resultados.xls en el que se almacena el análisis exportado

Utilizar el módulo de programabilidad ejecutando SPSS de forma que proporciona grandes ventajas a la hora de realizar el proyecto.

La primera y fundamental es que SPSS es un software completísimo y líder absoluto en su categoría, por lo que presta cualquier tipo de función estadística que se pueda necesitar, evitando así tener que desarrollarlas.

Otra gran ventaja es que estas funciones estadísticas que se han necesitado para realizar la aplicación pueden ser editadas con facilidad a nivel de código fuente, ya que la forma de comunicar el software con SPSS se basa en la concatenación de instrucciones separadas.

Otra gran ventaja es que al ejecutarse SPSS por debajo del usuario final de la aplicación se encuentra enfrente de una herramienta sencilla e intuitiva de manejar, y no ante un software de gran tamaño y cientos de opciones que puede imponer mucho respeto a cierta cantidad de potenciales usuarios que o bien por no estar habituados a trabajar con grandes programas o bien por malas experiencias anteriores piensen que no van a ser capaces de utilizarlo y entenderlo.

Por otro lado, también han existido inconvenientes a la hora de trabajar con este módulo. La más importante es la absoluta y total dependencia de las bibliotecas proporcionadas por SPSS.

A la hora de realizar cualquier operación que involucrara de una forma u otra trabajo con SPSS se necesita lanzar un comando de conexión entre el software y SPSS y por tanto se necesitaba que ese comando proporcionado funcionara perfectamente y realizara la función exacta para la que había sido preconcebido.

Al ser SPSS una compañía en constante evolución es muy común que se vayan descubriendo errores en su herramienta con sus correspondientes soluciones. Pero también ocurre que a veces estas mejoras se incluyen en versiones futuras cuya fecha de lanzamiento fue posterior o paralela a la realización de este proyecto, como ha ocurrido con los gráficos que se muestran en el proyecto, que son de elaboración propia y no autogenerados por SPSS. Y es que para cada versión de SPSS requiere su propio módulo de programabilidad, pero para el proyecto se debía trabajar en una versión fija, la que proporciona la Universidad Complutense y que se encuentra instalada en los laboratorios de la facultad.

Otro problema derivado de la dependencia de SPSS es la falta de documentación sobre éstos y la dificultad que entraña cambiar algunas de las opciones que generan por defecto. Así pues, estos comandos han sido un arma de doble filo, si bien se podía usar un mismo comando para realizar multitud de análisis distintos modificando dinámicamente algunas partes de su estructura, cambiar otras de las partes que generaba era problemático por no conocer si la opción que aparecía era única o se podía modificar para adaptarlo a las necesidades.

La función de este proyecto es cubrir unas necesidades muy específicas dentro de un tema ámbito bastante cerrado, como es el estadístico. Por lo que las alternativas a como se realizó el proyecto no eran demasiadas.

Por un lado se tenía la opción de trabajar con SAS, otro software, también bastante importante en el campo de la estadística, que ofrece la oportunidad de crear programas usando su propio lenguaje, para luego ser ejecutado por el propio programa. De esta idea hubo varias cosas que no convencieron. La primera era que el lenguaje que SAS proporciona era bastante limitado para construir un software muy visual e intuitivo, que era uno de los principales objetivos. El lenguaje de SAS sólo permite construir procedimientos o funciones que realicen cálculos para luego ser ejecutados por el propio SAS, por lo que además se debía buscar la forma de conectar lo realizado en SAS con otro framework que ofreciera mayores alternativas en lo que a lo visual se refiere.

Otra opción que en su momento se planteó muy seriamente fue la de construir el software entero, sin el apoyo de ningún programa estadístico y construir los procedimientos necesarios para realizar los análisis convenientes y devolver los resultados obtenidos. Pero al contar con esa gran ventaja que SPSS ofrecía también se desechó la opción.

Dado que estaba decidido que el proyecto se apoyaría en las funciones de programabilidad predefinidas en el módulo de SPSS, existían tres alternativas para la implementación en cuando a lenguajes de programación se refiere. La primera alternativa era usar alguno de los lenguajes soportados por el framework de .Net, otra opción era realizarlo con Python y la última, R.

En las primeras fases del proyecto, la opción con más fuerza es la de realizar una implementación en Python, influenciados sobre todo porque la mayor parte de la documentación y ejemplos ofrecidos por SPSS iban orientados a ese lenguaje. Además, también se consultó vía foro de la compañía a su ingeniero jefe, John Peck sobre la mejor forma de orientar el proyecto que teníamos entre manos. Su recomendación era también realizarlo en Python.

Python además ofrecía el atractivo de ser un lenguaje relativamente nuevo, en franca expansión y cada vez más popular entre los programadores debido a su gran potencia y a su sencillez. En este punto el grupo del proyecto se puso en busca de intérpretes de Python que nos ayudaran a realizar una interfaz amigable y que nos permitirán desarrollar con relativa rapidez todas las componentes necesarias. Sin embargo no se encontraron grandes soluciones para cumplir los requisitos que se deseaban para el proyecto y a eso se sumaba que tampoco se era capaz de conseguir buenos resultados de estudios implementando con Python. Estaba claro que se necesitaba un cambio de rumbo en la orientación del proyecto.

Aproximadamente en las mismas fechas en las que ocurrían estos contratiempos, SPSS liberaba gran cantidad de información de programabilidad para .Net, principalmente para Visual Basic. Pero una vez reconocidos los lenguajes incluidos en las tecnologías .Net, nos decantamos por C#.

C# es un lenguaje de programación orientado a objetos ideado y desarrollado por Microsoft. Su sintaxis deriva de la de C y C++ y su modelo de objetos es muy similar al de Java. Se podría decir que es el híbrido de Microsoft que aglutina características de los lenguajes orientados a objetos más importantes. Pese a ser un lenguaje que nunca se había usado, no fue difícil hacerse con sus fundamentos por los conocimientos que se poseían tanto de Java como de C++ y de C. A destacar de C#, que como Java gestiona memoria automáticamente, aunque también se puede realizar gestión de punteros a memoria como se haría con los lenguajes de C y C++. Como se ha explicado anteriormente las otras alternativas que había eran Python y R. De Python ya se ha hablado anteriormente en esta memoria, ya que fue el primer lenguaje con el que se pensó desarrollar el proyecto.

De la otra alternativa real que se tenía, lenguaje R, no se puede contar demasiados detalles, ya que es una opción que no se llegó a plantear seriamente a la hora del desarrollo del proyecto. Se puede decir de él, que es un lenguaje de programación y un entorno de desarrollo orientado plenamente al apartado estadístico y gráfico. R es un proyecto software libre basado en el lenguaje S y cada vez más popular en campos como la bioinformática, o en las matemáticas financieras. Y se puede utilizar tanto en sistemas Linux, como en Mac como en Windows.

No se ha podido indagar lo suficiente en este último lenguaje de programación, pero lo poco que se ha visto hace pensar que puede ser muy interesante para el desarrollo futuro de herramientas parecidas a la nuestra.

El entorno de desarrollo por el que se ha optado para la elaboración del proyecto ha sido Microsoft Visual Studio 2008, debido a que este software se encuentra plenamente adaptado al lenguaje C#, ya que ambos pertenecen a la misma compañía.

Además de las bibliotecas que se debían usar para conectar nuestro software con SPSS y que venían incluidas en el módulo de programabilidad, ha habido que usar otra biblioteca específica para realizar los gráficos que se generan en el proyecto.

La biblioteca escogida se llama WebChart y es de libre distribución, se puede descargar gratuitamente de internet. La inclusión y uso de esta biblioteca adicional viene condicionada por un error dentro del software de SPSS, y es que para la versión 17 del mismo, con la que se estaba trabajando, las funciones incluidas en el módulo de programabilidad no son capaces de generar los gráficos que se muestran si se realiza ese mismo estudio desde SPSS directamente.

Por ello hubo que investigar el apartado gráficos que tiene C# por defecto, y al no convencer en exceso se optó por la búsqueda y uso de alguna biblioteca especializada, entre las distintas opciones que se encontraron, como nPlot, se eligió WebChart, ya que ofrecía un interesante manual de usuario y ejemplos para poder practicar y mejorar nosotros mismos.

## Casos de uso

Se procede a la realización de un diagrama de casos de uso (*Main Use Case Diagram*). El objetivo es la creación de una representación gráfica del total de los actores y casos de uso del sistema (funcionalidades principales), incluyendo sus interacciones. En la figura 10 se puede observar el diagrama de casos de uso de la aplicación.

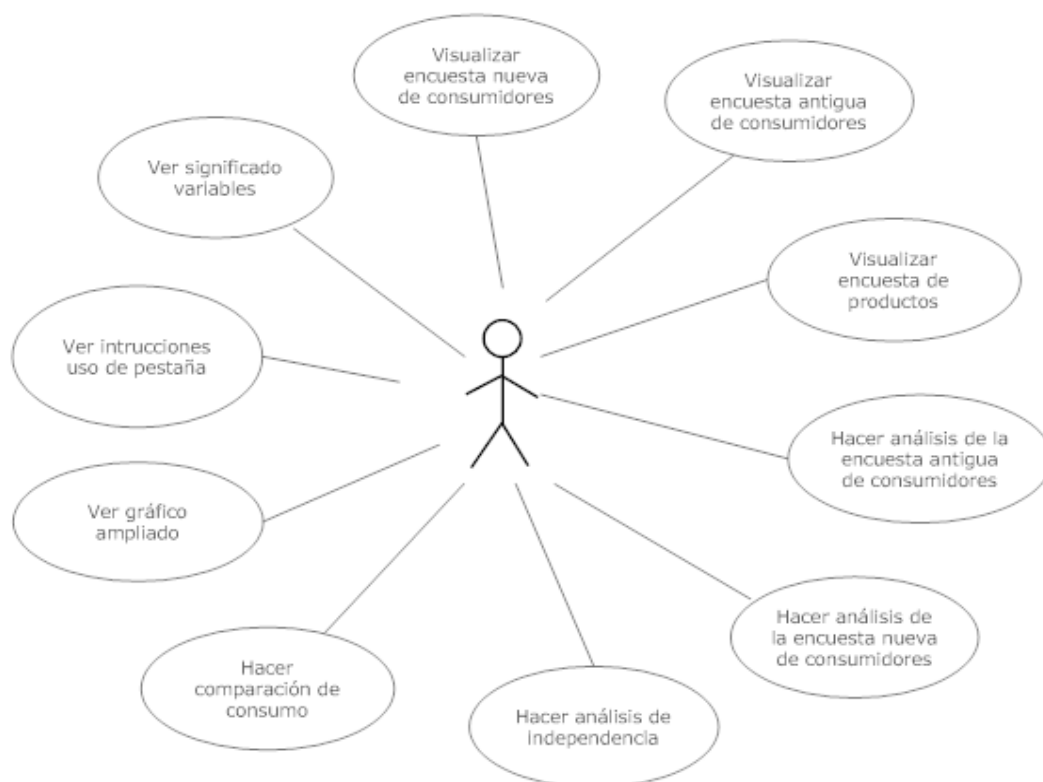


Figura 10. Diagrama de casos de uso de la aplicación

Como se puede apreciar, un usuario (único rol del sistema) podrá realizar las siguientes acciones:

- *Visualizar encuesta antigua de consumidores:* Se muestra el contenido de la base de datos de la encuesta antigua de consumidores.
- *Visualizar encuesta nueva de consumidores:* Se muestra el contenido de la base de datos de la encuesta nueva de consumidores.
- *Visualizar encuesta de productos:* Se muestra el contenido de las bases de datos de los productos de MDD.
- *Hacer análisis de la encuesta antigua de consumidores:* Se hacen diferentes análisis con los datos de la encuesta antigua de consumidores. El usuario puede elegir entre varios estudios y opciones para los estudios. Se muestra un gráfico con los resultados y se genera un informe en formato Excel.
- *Hacer análisis de la encuesta nueva de consumidores:* Se hacen diferentes análisis con los datos de la encuesta nueva de consumidores. El usuario puede elegir entre varios estudios y opciones para los estudios. Se muestra un gráfico con los resultados y se genera un informe en formato Excel.
- *Hacer análisis de independencia:* Se hace un análisis de independencia entre dos variables de las encuestas de los consumidores. El usuario debe elegir sobre qué encuesta se realiza el estudio y las variables sobre las cuales se hará. Se genera un informe con los resultados en formato Excel.
- *Hacer comparación de consumo:* Se realiza una comparación de consumo por rango de edad y ola. El usuario debe elegir la ola sobre la que se realiza la comparación y la ponderación de las diferentes factores que se tienen en cuenta.
- *Ver grafico ampliado:* Se muestra el gráfico señalado por el usuario ampliado con una leyenda detallada.
- *Ver instrucciones de uso de pestaña:* Se muestra las instrucciones de uso de la pestaña seleccionada.
- *Ver significado de pestañas:* Se explica el significado de las diferentes variables de las BB.DD.

### 3.3. Diseño y modelado de las Bases de Datos

Los datos de entrada que se van a usar para realizar los diferentes análisis y que rigen todo el proyecto son unas Bases de Datos (a partir de ahora BB.DD) en la que se encuentran almacenadas tanto las encuestas a los consumidores en la puerta de los establecimientos, como la información de los productos de MDD.

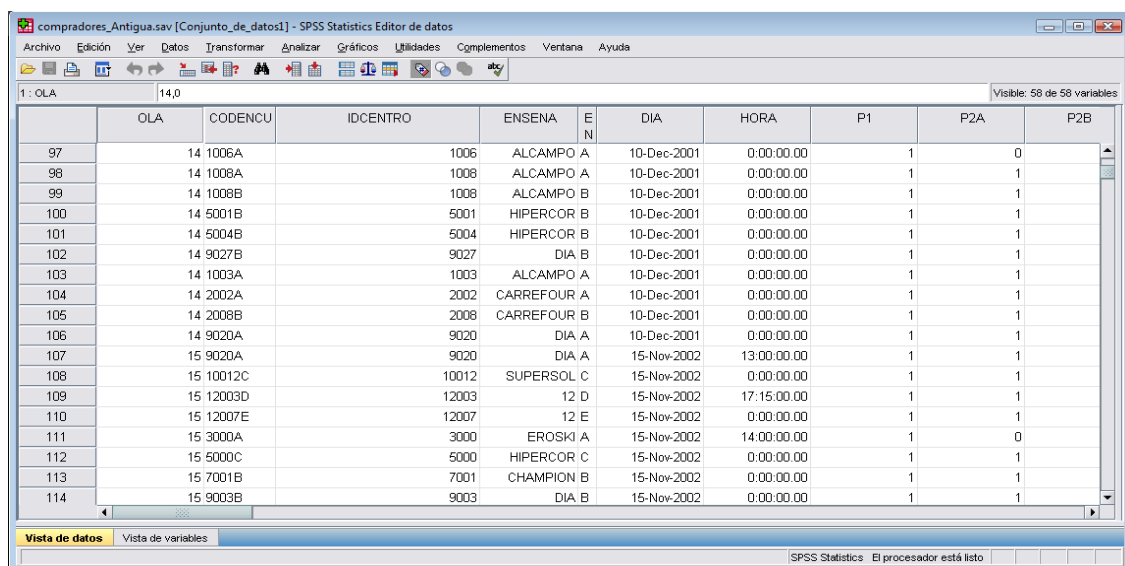
Las diferentes BB.DD que se usan están en formato SPSS, es muy fácil de introducir los datos y de importar o exportar los datos a otros formatos. Además permite directamente poder hacer los diferentes estudios sobre los datos, lo que facilita el trabajo a la hora de hacer dichos estudios.

A continuación se explica de una forma más detallada las diferentes BB.DD usadas a lo largo del proyecto.

### Compradores \_ antigua

Dentro de esta Base de Datos (a partir de ahora B.D) se encuentra los datos de las encuestas de compradores de las olas 14 a 19. Está formada por unas cuatro mil encuestas, cada una de ella con unas 56 variables; tanto datos de la encuesta y consumidor, como respuestas a las preguntas de las encuestas. En la figura 11 podemos observar la B.D de encuestas antiguas a consumidores en SPSS.

Esta B.D se usa en el programa en la visualización de las encuestas de compradores antigua, en la realización de estudios sobre ella, en la realización de estudios de independencia y, por último, en el estudio de comparación de consumo.



|     | OLA | CODENCU | IDCENTRO | ENSEÑA    | E N | DIA         | HORA        | P1 | P2A | P2B |
|-----|-----|---------|----------|-----------|-----|-------------|-------------|----|-----|-----|
| 97  | 14  | 1006A   | 1006     | ALCAMPO   | A   | 10-Dec-2001 | 0:00:00.00  | 1  |     | 0   |
| 98  | 14  | 1008A   | 1008     | ALCAMPO   | A   | 10-Dec-2001 | 0:00:00.00  | 1  |     | 1   |
| 99  | 14  | 1008B   | 1008     | ALCAMPO   | B   | 10-Dec-2001 | 0:00:00.00  | 1  |     | 1   |
| 100 | 14  | 5001B   | 5001     | HIPERCOR  | B   | 10-Dec-2001 | 0:00:00.00  | 1  |     | 1   |
| 101 | 14  | 5004B   | 5004     | HIPERCOR  | B   | 10-Dec-2001 | 0:00:00.00  | 1  |     | 1   |
| 102 | 14  | 9027B   | 9027     | DIA       | B   | 10-Dec-2001 | 0:00:00.00  | 1  |     | 1   |
| 103 | 14  | 1003A   | 1003     | ALCAMPO   | A   | 10-Dec-2001 | 0:00:00.00  | 1  |     | 1   |
| 104 | 14  | 2002A   | 2002     | CARREFOUR | A   | 10-Dec-2001 | 0:00:00.00  | 1  |     | 1   |
| 105 | 14  | 2008B   | 2008     | CARREFOUR | B   | 10-Dec-2001 | 0:00:00.00  | 1  |     | 1   |
| 106 | 14  | 9020A   | 9020     | DIA       | A   | 10-Dec-2001 | 0:00:00.00  | 1  |     | 1   |
| 107 | 15  | 9020A   | 9020     | DIA       | A   | 15-Nov-2002 | 13:00:00.00 | 1  |     | 1   |
| 108 | 15  | 10012C  | 10012    | SUPERSOL  | C   | 15-Nov-2002 | 0:00:00.00  | 1  |     | 1   |
| 109 | 15  | 12003D  | 12003    | 12 D      |     | 15-Nov-2002 | 17:15:00.00 | 1  |     | 1   |
| 110 | 15  | 12007E  | 12007    | 12 E      |     | 15-Nov-2002 | 0:00:00.00  | 1  |     | 1   |
| 111 | 15  | 3000A   | 3000     | EROSKI    | A   | 15-Nov-2002 | 14:00:00.00 | 1  |     | 0   |
| 112 | 15  | 5000C   | 5000     | HIPERCOR  | C   | 15-Nov-2002 | 0:00:00.00  | 1  |     | 1   |
| 113 | 15  | 7001B   | 7001     | CHAMPION  | B   | 15-Nov-2002 | 0:00:00.00  | 1  |     | 1   |
| 114 | 15  | 9003B   | 9003     | DIA       | B   | 15-Nov-2002 | 0:00:00.00  | 1  |     | 1   |

Figura11. B.D compradores antigua

La B.D están formadas por las variables que se muestran en la tabla 1:

|    |          |   |
|----|----------|---|
| 1  | OLA      | Ola de la que es la encuesta (14, 15, 16, 17, 18 , 19)  |
| 2  | CODENCU  | Número de centro y tipo de encuesta   |
| 3  | IDCENTRO | Número del centro   |
| 4  | ENSEÑA   | Nombre del supermercado   |
| 5  | ENCUESTA | Tipo de encuesta realizada (de la A a la L)   |
| 6  | DIA      | Día en el que se hizo la encuesta   |
| 7  | HORA     | Hora en la que se hizo la encuesta  |
| 8  | P1       | P1.- ¿Conoce, aunque no los haya comprado nunca, productos de alimentación que tienen la marca del establecimiento que los vende? |
| 9  | P2A      | P2 ¿Conoce usted productos de MDD de tipoõ .? a.- Lácteos y asimilados  |
| 10 | P2B      | P2 ¿Compra usted productos de MDD de tipoõ ? a.- Lácteos y asimilados   |



|    |      |   |
|----|------|---|
| 11 | P3A  | P3 ¿Conoce usted productos de MDD de tipo ñ .? b.- Aceites  |
| 12 | P3B  | P3 ¿Compra usted productos de MDD de tipo ñ ? b.- Aceites   |
| 13 | P4A  | P4 ¿Conoce usted productos de MDD de tipo ñ .? c.- Cafés  |
| 14 | P4B  | P4 ¿Compra usted productos de MDD de tipo ñ ? c.- Cafés   |
| 15 | P5A  | P5 ¿Conoce usted productos de MDD de tipo ñ .? d.- Chocolates   |
| 16 | P5B  | P5 ¿Compra usted productos de MDD de tipo ñ ? d.- Chocolates  |
| 17 | P6A  | P6 ¿Conoce usted productos de MDD de tipo ñ .? e.- Mermeladas y miel  |
| 18 | P6B  | P6 ¿Compra usted productos de MDD de tipo ñ ? e.- Mermeladas y miel   |
| 19 | P7A  | P7 ¿Conoce usted productos de MDD de tipo ñ .? f.- Galletas y pan   |
| 20 | P7B  | P7 ¿Compra usted productos de MDD de tipo ñ ? f.- Galletas y pan  |
| 21 | P8A  | P8 ¿Conoce usted productos de MDD de tipo ñ .? g.- Legumbres y varios   |
| 22 | P8B  | P8 ¿Compra usted productos de MDD de tipo ñ ? g.- Legumbres y varios  |
| 23 | P9A  | P9 ¿Conoce usted productos de MDD de tipo ñ .? h.- Conservas de pescado   |
| 24 | P9B  | P9 ¿Compra usted productos de MDD de tipo ñ ? h.- Conservas de pescado  |
| 25 | P10A | P10 ¿Conoce usted productos de MDD de tipo ñ .? i.- Conservas vegetales   |
| 26 | P10B | P10 ¿Compra usted productos de MDD de tipo ñ ? i.- Conservas vegetales  |
| 27 | P11A | P11 ¿Conoce usted productos de MDD de tipo ñ .? j.- Pastas  |
| 28 | P11B | P11 ¿Compra usted productos de MDD de tipo ñ ? j.- Pastas   |
| 29 | P12A | P12 ¿Conoce usted productos de MDD de tipo ñ .? k.- Bebidas   |
| 30 | P12B | P12 ¿Compra usted productos de MDD de tipo ñ ? k.- Bebidas  |
| 31 | P13A | P13 ¿Conoce usted productos de MDD de tipo ñ .? l.- Turrónes  |
| 32 | P13B | P13 ¿Compra usted productos de MDD de tipo ñ ? l.- Turrónes   |
| 33 | P14A | P14 ¿Cuáles son las razones para comprar en este supermercado? Precios bajos  |
| 34 | P14B | P14 ¿Cuáles son las razones para comprar en este supermercado? Distancia  |
| 35 | P14C | P14 ¿Cuáles son las razones para comprar en este supermercado? Variedad de productos  |
| 36 | P14D | P14 ¿Cuáles son las razones para comprar en este supermercado? Buena calidad  |
| 37 | P14E | P14 ¿Cuáles son las razones para comprar en este supermercado? Garantía del establecimiento                                   |
| 38 | P14F | P14 ¿Cuáles son las razones para comprar en este supermercado? Otras (publicidad, degustaciones)                              |
| 39 | P15  | P15.- En relación con el año pasado, ¿usted compra más, menos o igual en el supermercado?                                     |
| 40 | P16  | P16.- ¿Compra productos con Marcas del Distribuidor más, menos o igual que el año pasado?                                     |
| 41 | P17  | P17 Importancia que tiene a la hora de su compra como es...a.- Precio   |
| 42 | P18  | P18 Importancia que tiene a la hora de su compra como es...b.- Envase   |
| 43 | P19  | P19 Importancia que tiene a la hora de su compra como es...c.- Calidad  |
| 44 | P20  | P20 Importancia que tiene a la hora de su compra como es...d.- Garantía   |
| 45 | P21  | P21 ¿Cree usted que las Marcas de Distribuidor son más baratas por ser más barato el envase e invertir menos en publicidad?   |
| 46 | P22  | P22 ¿Cree usted que las Marcas de Distribuidor y las Marcas de Fabricante son elaboradas por los mismos fabricantes?          |
| 47 | P23  | P23 ¿Cree usted que las Marcas de Distribuidor son tan buenas como las Marcas de Fabricante                                   |
| 48 | P24  | P24.- ¿Qué influencia tienen las promociones especiales para que se decida a probar las Marcas de Distribución?               |
| 49 | P25  | P25.- ¿Qué influencia tienen la garantía de satisfacción de la tienda para que se decida a probar las Marcas de Distribución? |
| 50 | P26  | P26 Datos de clasificación Edad   |

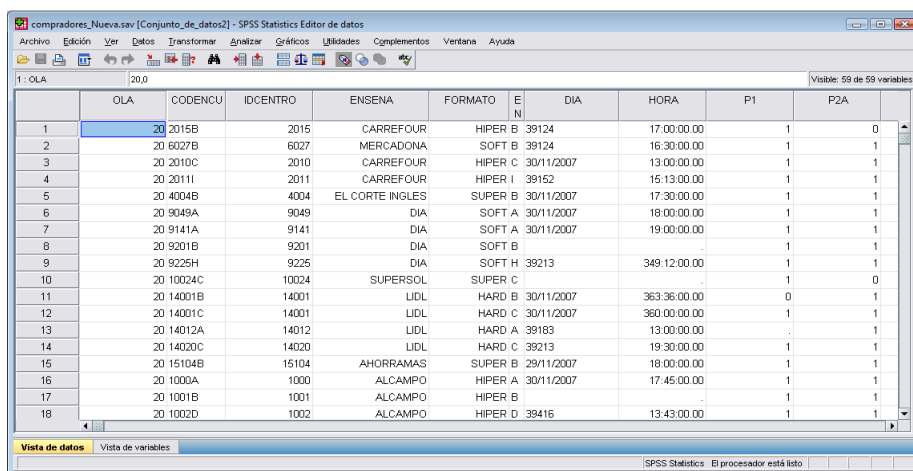
|    |     |  |
|----|-----|--|
| 51 | P27 | P27 Datos de clasificación Status  |
| 52 | P28 | P28 Datos de clasificación Número de Personas que viven en hogar         |
| 53 | P29 | P29 Datos de clasificación Distancia al lugar de residencia              |
| 54 | P30 | P30 Datos de clasificación Frecuencia de la compra en general            |
| 55 | P31 | P31 Datos de clasificación Frecuencia de la compra en el establecimiento |
| 56 | P32 | P32 Datos de clasificación Sexo  |

**Tabla 1.** Tabla con las variables de la B.D Compradores\_antigua y su significado

## Compradores \_nueva

Dentro de esta B.D se encuentra los datos de las encuestas de compradores de las olas a partir de la 20. Cada encuesta está formada por unas 57 variables; tanto datos de la encuesta y consumidor, como respuestas a las preguntas de las encuestas. En la figura 12 podemos observar la B.D de encuestas nuevas a consumidores en SPSS.

Esta B.D se usa en el programa en la visualización de las encuestas de compradores nueva, en la realización de estudios sobre ella y, por último, en la realización de estudios de independencia.



|    | OLA | CODENCU | IDCENTRO | ENSEÑA          | FORMATO | E N        | DIA | HORA         | P1 | P2A |
|----|-----|---------|----------|-----------------|---------|------------|-----|--------------|----|-----|
| 1  | 20  | 2015B   | 2015     | CARREFOUR       | HIPER B | 39124      |     | 17:00:00.00  | 1  | 0   |
| 2  | 20  | 6027B   | 6027     | MERCADONA       | SOFT B  | 39124      |     | 16:30:00.00  | 1  | 1   |
| 3  | 20  | 2010C   | 2010     | CARREFOUR       | HIPER C | 30/11/2007 |     | 13:00:00.00  | 1  | 1   |
| 4  | 20  | 2011I   | 2011     | CARREFOUR       | HIPER I | 39152      |     | 15:13:00.00  | 1  | 1   |
| 5  | 20  | 4004B   | 4004     | EL CORTE INGLES | SUPER B | 30/11/2007 |     | 17:30:00.00  | 1  | 1   |
| 6  | 20  | 9049A   | 9049     | DIA             | SOFT A  | 30/11/2007 |     | 18:00:00.00  | 1  | 1   |
| 7  | 20  | 9141A   | 9141     | DIA             | SOFT A  | 30/11/2007 |     | 19:00:00.00  | 1  | 1   |
| 8  | 20  | 9201B   | 9201     | DIA             | SOFT B  |            |     |              | 1  | 1   |
| 9  | 20  | 9225H   | 9225     | DIA             | SOFT H  | 39213      |     | 349:12:00.00 | 1  | 1   |
| 10 | 20  | 10024C  | 10024    | SUPERSOL        | SUPER C |            |     |              | 1  | 0   |
| 11 | 20  | 14001B  | 14001    | LIDL            | HARD B  | 30/11/2007 |     | 363:36:00.00 | 0  | 1   |
| 12 | 20  | 14001C  | 14001    | LIDL            | HARD C  | 30/11/2007 |     | 360:00:00.00 | 1  | 1   |
| 13 | 20  | 14012A  | 14012    | LIDL            | HARD A  | 39183      |     | 13:00:00.00  |    | 1   |
| 14 | 20  | 14020C  | 14020    | LIDL            | HARD C  | 39213      |     | 19:30:00.00  | 1  | 1   |
| 15 | 20  | 15104B  | 15104    | AHORRAMAS       | SUPER B | 29/11/2007 |     | 18:00:00.00  | 1  | 1   |
| 16 | 20  | 1000A   | 1000     | ALCAMPO         | HIPER A | 30/11/2007 |     | 17:45:00.00  | 1  | 1   |
| 17 | 20  | 1001B   | 1001     | ALCAMPO         | HIPER B |            |     |              | 1  | 1   |
| 18 | 20  | 1002D   | 1002     | ALCAMPO         | HIPER D | 39416      |     | 13:43:00.00  | 1  | 1   |

**Figura12.** B.D compradores nueva

La B.D están formadas por las variables que se muestran en la tabla 2:

|   |          |   |
|---|----------|---|
| 1 | OLA      | Ola de la que es la encuesta  |
| 2 | CODENCU  | Número de centro y tipo de encuesta   |
| 3 | IDCENTRO | Número del centro   |
| 4 | ENSEÑA   | Nombre del supermercado   |
| 5 | FORMATO  | Formato del supermercado (HIPER, SUPER, SUPER SOFT, SOFT, HARD)   |
| 6 | ENCUESTA | Tipo de encuesta realizada (de la A a la L)   |
| 7 | DIA      | Día en el que se hizo la encuesta   |
| 8 | HORA     | Hora en la que se hizo la encuesta  |
| 9 | P1       | P1.- ¿Conoce, aunque no los haya comprado nunca, productos de alimentación que tienen la marca del establecimiento que los vende? |

|    |      |   |
|----|------|---|
| 10 | P2A  | P2 ¿Conoce usted productos de MDD de tipo ñ .? a.- Lácteos y asimilados   |
| 11 | P2B  | P2 ¿Compra usted productos de MDD de tipo ñ ? a.- Lácteos y asimilados  |
| 12 | P3A  | P3 ¿Conoce usted productos de MDD de tipo ñ .? b.- Aceites  |
| 13 | P3B  | P3 ¿Compra usted productos de MDD de tipo ñ ? b.- Aceites   |
| 14 | P4A  | P4 ¿Conoce usted productos de MDD de tipo ñ .? c.- Cafés  |
| 15 | P4B  | P4 ¿Compra usted productos de MDD de tipo ñ ? c.- Cafés   |
| 16 | P5A  | P5 ¿Conoce usted productos de MDD de tipo ñ .? d.- Chocolates   |
| 17 | P5B  | P5 ¿Compra usted productos de MDD de tipo ñ ? d.- Chocolates  |
| 18 | P6A  | P6 ¿Conoce usted productos de MDD de tipo ñ .? e.- Mermeladas y miel  |
| 19 | P6B  | P6 ¿Compra usted productos de MDD de tipo ñ ? e.- Mermeladas y miel   |
| 20 | P7A  | P7 ¿Conoce usted productos de MDD de tipo ñ .? f.- Galletas y pan   |
| 21 | P7B  | P7 ¿Compra usted productos de MDD de tipo ñ ? f.- Galletas y pan  |
| 22 | P8A  | P8 ¿Conoce usted productos de MDD de tipo ñ .? g.- Legumbres y varios   |
| 23 | P8B  | P8 ¿Compra usted productos de MDD de tipo ñ ? g.- Legumbres y varios  |
| 24 | P9A  | P9 ¿Conoce usted productos de MDD de tipo ñ .? h.- Conservas de pescado   |
| 25 | P9B  | P9 ¿Compra usted productos de MDD de tipo ñ ? h.- Conservas de pescado  |
| 26 | P10A | P10 ¿Conoce usted productos de MDD de tipo ñ .? i.- Conservas vegetales   |
| 27 | P10B | P10 ¿Compra usted productos de MDD de tipo ñ ? i.- Conservas vegetales  |
| 28 | P11A | P11 ¿Conoce usted productos de MDD de tipo ñ .? j.- Pastas  |
| 29 | P11B | P11 ¿Compra usted productos de MDD de tipo ñ ? j.- Pastas   |
| 30 | P12A | P12 ¿Conoce usted productos de MDD de tipo ñ .? k.- Bebidas   |
| 31 | P12B | P12 ¿Compra usted productos de MDD de tipo ñ ? k.- Bebidas  |
| 32 | P13A | P13 ¿Conoce usted productos de MDD de tipo ñ .? l.- Turrónes  |
| 33 | P13B | P13 ¿Compra usted productos de MDD de tipo ñ ? l.- Turrónes   |
| 34 | P14A | P14 ¿Cuáles son las razones para comprar en este supermercado? Precios bajos  |
| 35 | P14B | P14 ¿Cuáles son las razones para comprar en este supermercado? Distancia  |
| 36 | P14C | P14 ¿Cuáles son las razones para comprar en este supermercado? Variedad de productos  |
| 37 | P14D | P14 ¿Cuáles son las razones para comprar en este supermercado? Buena calidad  |
| 38 | P14E | P14 ¿Cuáles son las razones para comprar en este supermercado? Garantía del establecimiento                                 |
| 39 | P14F | P14 ¿Cuáles son las razones para comprar en este supermercado? Otras (publicidad, degustaciones)                            |
| 40 | P15  | P15.- En relación con el año pasado, ¿usted compra más, menos o igual en el supermercado?                                   |
| 41 | P16  | P16.- ¿Compra productos con Marcas del Distribuidor más, menos o igual que el año pasado?                                   |
| 42 | P17  | P17 Importancia que tiene a la hora de su compra como es...a.- Precio   |
| 43 | P18  | P18 Importancia que tiene a la hora de su compra como es...b.- Envase   |
| 44 | P19  | P19 Importancia que tiene a la hora de su compra como es...c.- Calidad  |
| 45 | P20  | P20 Importancia que tiene a la hora de su compra como es...d.- Garantía   |
| 46 | P21  | P21 ¿Cree usted que las Marcas de Distribuidor son más baratas por ser más barato el envase e invertir menos en publicidad? |
| 47 | P22  | P22 ¿Cree usted que las Marcas de Distribuidor y las Marcas de Fabricante son elaboradas por los mismos fabricantes?        |
| 48 | P23  | P23 ¿Cree usted que las Marcas de Distribuidor son tan buenas como las Marcas de Fabricante                                 |
| 49 | P24  | P24.- ¿Qué influencia tienen las promociones especiales para que se decida a probar las Marcas de Distribución?             |

|    |     |   |
|----|-----|---|
| 50 | P25 | P25.- ¿Qué influencia tienen la garantía de satisfacción de la tienda para que se decida a probar las Marcas de Distribución? |
| 51 | P26 | P26 Datos de clasificación Edad   |
| 52 | P27 | P27 Datos de clasificación Status   |
| 53 | P28 | P28 Datos de clasificación Número de Personas que viven en hogar  |
| 54 | P29 | P29 Datos de clasificación Distancia al lugar de residencia   |
| 55 | P30 | P30 Datos de clasificación Frecuencia de la compra en general   |
| 56 | P31 | P31 Datos de clasificación Frecuencia de la compra en el establecimiento  |
| 57 | P32 | P32 Datos de clasificación Sexo   |

**Tabla 2.** Tabla con las variables de la B.D Compradores\_nueva y su significado

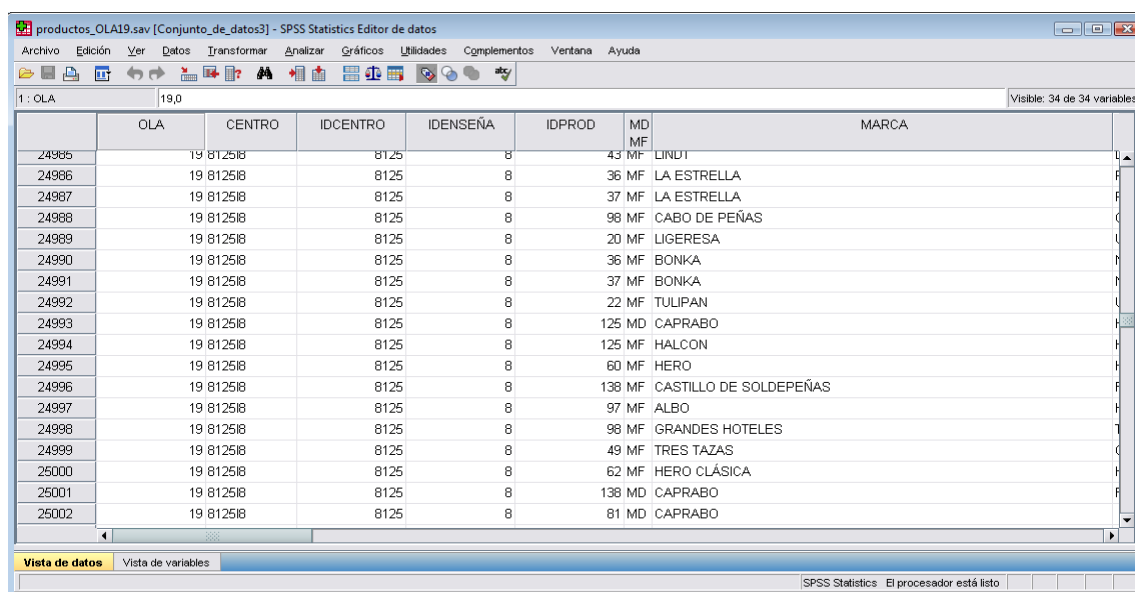
La diferencia que existe con respecto a la B.DD consumidores \_ antigua es que se le ha añadido una nueva variable denominada Formato, en donde se explica el formato del supermercado en el que se ha hecho la encuesta. Pero la mayor diferencia con respecto a la B.DD consumidores \_ antigua es la puntuación que se le da a las preguntas, han pasado de una puntuación de 1 a 5, en el caso de consumidores\_antigua, a una puntuación de 1 a 7 en el caso de consumidores\_nueva.

A día de hoy en esta B.DD se encuentran los datos de la Ola 20, pero puede ser ampliado con otras olas, ola 21, ola22 etc.; siempre y cuando los datos tengan el formato de la tabla anterior. Esto permite que se pueda seguir realizando estudios con una mayor cantidad de datos, lo que permite resultados más fiables y que la aplicación se pueda utilizar con nuevas olas.

### Productos\_ OlaXX

Las BB.DD de los productos están divididas por Olas, esto es debido a que el tamaño de los datos de cada Ola es muy grande, por lo que se tiene que guardar de una en una.

Tenemos siete BB.DD de los productos que son las siete Olas de las que se dispone, de la 14 a la 20; cada una de ellas denominada productos\_OLANºOla (productos\_OLA15). Cada B.D está formada por más de veinte mil datos de productos diferentes, en algunas olas superan los treinta mil. Cada fila de las BB.DD está formada por 34 variables, en las que se almacena información del producto, situación, precio, marca, comparación con la marca líder etc. En la figura 13 podemos observar la B.D de productos de la ola 19 en SPSS.



|       | OLA | CENTRO | IDCENTRO | IDENSEÑA | IDPROD | MDMF | MARCA                  |
|-------|-----|--------|----------|----------|--------|------|------------------------|
| 24985 | 19  | 812518 | 8125     | 8        | 43     | MF   | LINDT                  |
| 24986 | 19  | 812518 | 8125     | 8        | 36     | MF   | LA ESTRELLA            |
| 24987 | 19  | 812518 | 8125     | 8        | 37     | MF   | LA ESTRELLA            |
| 24988 | 19  | 812518 | 8125     | 8        | 98     | MF   | CABO DE PEÑAS          |
| 24989 | 19  | 812518 | 8125     | 8        | 20     | MF   | LIGERESA               |
| 24990 | 19  | 812518 | 8125     | 8        | 36     | MF   | BONKA                  |
| 24991 | 19  | 812518 | 8125     | 8        | 37     | MF   | BONKA                  |
| 24992 | 19  | 812518 | 8125     | 8        | 22     | MF   | TULIPAN                |
| 24993 | 19  | 812518 | 8125     | 8        | 125    | MD   | CAPRABO                |
| 24994 | 19  | 812518 | 8125     | 8        | 125    | MF   | HALCON                 |
| 24995 | 19  | 812518 | 8125     | 8        | 60     | MF   | HERO                   |
| 24996 | 19  | 812518 | 8125     | 8        | 138    | MF   | CASTILLO DE SOLDEPEÑAS |
| 24997 | 19  | 812518 | 8125     | 8        | 97     | MF   | ALBO                   |
| 24998 | 19  | 812518 | 8125     | 8        | 98     | MF   | GRANDES HOTELES        |
| 24999 | 19  | 812518 | 8125     | 8        | 49     | MF   | TRES TAZAS             |
| 25000 | 19  | 812518 | 8125     | 8        | 62     | MF   | HERO CLÁSICA           |
| 25001 | 19  | 812518 | 8125     | 8        | 138    | MD   | CAPRABO                |
| 25002 | 19  | 812518 | 8125     | 8        | 81     | MD   | CAPRABO                |

Figura13. B.D productos\_Ola19

Las BB.DD están formadas por las variables que se muestran en la tabla 3:

|    |            |                                   |
|----|------------|-----------------------------------|
| 1  | OLA        | Número de Ola                     |
| 2  | CENTRO     | Nombre del centro                 |
| 3  | IDCENTRO   | ID del centro                     |
| 4  | IDENSEÑA   | ID de la enseña                   |
| 5  | IDPROD     | ID del producto                   |
| 6  | MDMF       | Si es una MF o una MDD            |
| 7  | MARCA      | Marca del Producto                |
| 8  | FABRICANTE | Fabricante del producto           |
| 9  | TTU        | Capacidad del producto            |
| 10 | PCRU       | Capacidad Real de producto        |
| 11 | NUR        | Número de unidades en el producto |
| 12 | PRR        | Precio real                       |
| 13 | PRU        | Precio por unidad                 |
| 14 | LIDER      | Nombre del líder                  |

|    |             |   |
|----|-------------|---|
| 15 | MQL         | Envase mejor que la del líder                       |
| 16 | IQL         | Envase igual que la del líder                       |
| 17 | PQL         | Envase peor que la del líder                        |
| 18 | ESPM13      | Espacio del producto en las estanterías mayor a 1/3 |
| 19 | ESP13       | Espacio del producto en las estanterías menor a 1/3 |
| 20 | ESPI13      | Espacio del producto en las estanterías igual a 1/3 |
| 21 | TECHO       | Productos situado a la altura del techo             |
| 22 | OJOS        | Productos situado a la altura de los ojos           |
| 23 | MANOS       | Productos situado a la altura de las manos          |
| 24 | SUELO       | Productos situado a la altura del suelo             |
| 25 | cabecera    | Producto situado en la cabecera del pasillo         |
| 26 | isla        | Producto situado en la isla del pasillo             |
| 27 | chimenea    | Producto situado en la chimenea del pasillo         |
| 28 | otros       | Producto situado otro sitio del pasillo             |
| 29 | Folleto     | Existe folletos del producto                        |
| 30 | Publicidad  | Existe publicidad en el centro del producto         |
| 31 | Cartel      | Existen carteles del producto                       |
| 32 | degustación | Hay degustaciones del producto                      |
| 33 | Otras       | Hay otro tipo de promoción del producto             |
| 34 | SINPROMO    | No hay promoción del producto                       |

**Tabla 3.** Tabla con las variables de la B.D productos\_OlaXX y su significado

## 3.4. Producto Final

Hasta el momento se ha hablado en la memoria de muchos conceptos teóricos, necesidades del proyecto, tecnologías empleadas e incluso de las bases de datos utilizadas, pero se ha dicho poco o nada de las partes de la que está formado el proyecto.

Con el fin de colaborar en el estudio mediante una aplicación informática, el departamento de Arquitectura de Computadores de la Facultad de Informática de la Universidad Complutense ha realizado este proyecto dentro del programa Sistemas Informáticos correspondiente al proyecto Fin de Carrera.

La herramienta software que se ha desarrollado y cuyos resultados se muestran en esta memoria, tiene, a grandes rasgos, algunas partes destacables:

- **Visualización de los datos:** se puede visualizar los datos referidos a las OLAS de consumidores, sobre las que se centra el trabajo, y también las de productos almacenados en Bases de Datos.
- **Realización de análisis:** son dos pestañas, en las que se realizarán los estudios con respecto al comportamiento de las respuestas de los consumidores en la base de datos
- **Análisis de independencia:** se realiza un estudio de independencia entre dos de las variables de la base de datos.
- **Comparación de consumo usando métodos de decisión multicriterio:** Utilizando técnicas de soft computing, se realizarán estudios en los que se podrá determinar que rango poblacional es más propenso a consumir marca blanca
- **Gráficos:** pestaña que muestra los gráficos con detalle, de los estudios de análisis.
- **Ayuda:** se muestra un resumen detallado de las opciones de cada pestaña, y también el significado de las variables de la base de datos.

### 3.4.1. Módulo 1: Visualización de los datos

Las dos primeras opciones de la aplicación consisten en visualizar las diferentes BB.DD que se usan. La visualización de los datos de las encuestas a los consumidores y de los productos permite que el usuario pueda ver con qué datos se hacen los diferentes estudios, dándole una mayor tranquilidad y seguridad en los resultados de los estudios.

La visualización de los diferentes datos se divide en dos pestañas, la primera, denominada *View Costumer*, permite al usuario ver las ver los datos de las encuestas. Estos datos se dividen entre encuesta antigua y encuesta nueva. El usuario debe de elegir cuál de las dos encuestas quiere ver. En el caso de la encuesta antigua se puede elegir las olas que quiere ver, entre la ola 14 y la 19. Una vez seleccionado el tipo de encuesta y, si procede, las olas al pulsar sobre el botón *View* se mostraran los datos.

En la figura 14 se muestra, como ejemplo, la visualización de las olas 14,15 y 19

| OLA | CODENCU | IDCENTRO | ENSENA | ENCUESTA | DIA         | HORA  | P1 | P2A | P2B | P |
|-----|---------|----------|--------|----------|-------------|-------|----|-----|-----|---|
| 15  | 8002C   | 8002     | 8      | C        | 13256697600 | 0     | 1  | 1   | 0   | 1 |
| 15  | 8015G   | 8015     | 8      | G        | 13256697600 | 0     | 1  | 1   | 1   | 0 |
| 15  | 9009C   | 9009     | 9      | C        | 13256697600 | 0     | 1  | 1   | 1   | 1 |
| 15  | 9009F   | 9009     | 9      | F        | 13256697600 | 0     | 1  | 1   | 1   | 1 |
| 15  | 9017D   | 9017     | 9      | D        | 13256697600 | 0     | 1  | 1   | 1   | 1 |
| 15  | 9031D   | 9031     | 9      | D        | 13256697600 | 36000 | 1  | 1   | 1   | 1 |
| 15  | 9052A   | 9052     | 9      | A        | 13256697600 | 39600 | 1  | 1   | 1   | 1 |
| 15  | 9057B   | 9057     | 9      | B        | 13256697600 | 0     | 1  | 1   | 1   | 1 |
| 19  | 1001C   | 1001     | 1      | C        | 13383446400 | 48300 | 1  | 1   | 1   | 1 |
| 19  | 1001J   | 1001     | 1      | J        | 13384310400 | 48600 | 1  | 0   | 1   | 0 |
| 19  | 1007C   | 1007     | 1      | C        | 13385001600 | 67200 | 1  | 1   | 1   | 0 |
| 19  | 13005A  | 13005    | 13     | A        | 13383964800 | 63000 | 1  | 1   | 0   | 1 |
| 19  | 13005C  | 13005    | 13     | C        | 13383964800 | 63900 | 1  | 1   | 1   | 0 |
| 19  | 15007C  | 15007    | 15     | C        | 13384396800 | 50700 | 0  | 0   | 0   | 0 |
| 19  | 15017C  | 15017    | 15     | C        | 13384310400 | 39900 | 1  | 1   | 1   | 1 |
| 19  | 15099B  | 15099    | 15     | B        | 13384224000 | 64800 | 1  | 1   | 0   | 1 |
| 19  | 3001F   | 3001     | 3      | F        | 13382928000 | 0     | 1  | 1   | 1   | 0 |
| 19  | 33012A  | 33012    | 18     | A        | 13384396800 | 61200 | 1  | 0   | 1   | 0 |

☒ WAVE 14  
☒ WAVE 15  
☐ WAVE 16  
☐ WAVE 17  
☐ WAVE 18  
☒ WAVE 19

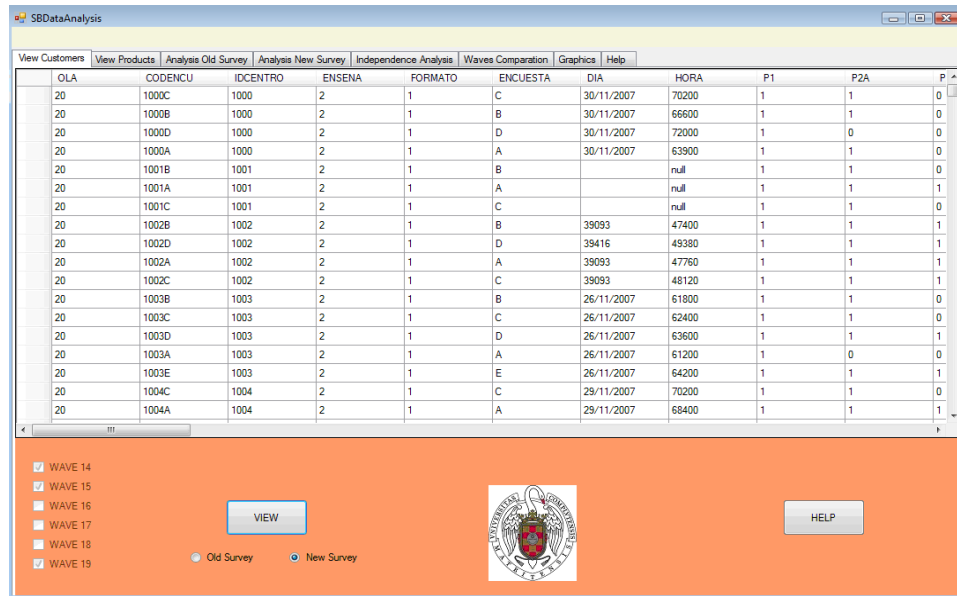
☐ Old Survey ☐ New Survey

**Figura 14.** Se muestra los datos de las encuestas de las Olas 14,15 y 19

En el caso de seleccionar la encuesta nueva se mostrará todos los datos que hay dentro de la B.D. Esto permite que se pueda añadir más datos a la B.D, de olas nuevas y que sea posible ver estos nuevos datos añadidos.



En la figura 15 se muestra la visualización de la B.D de encuestas nuevas a consumidores.

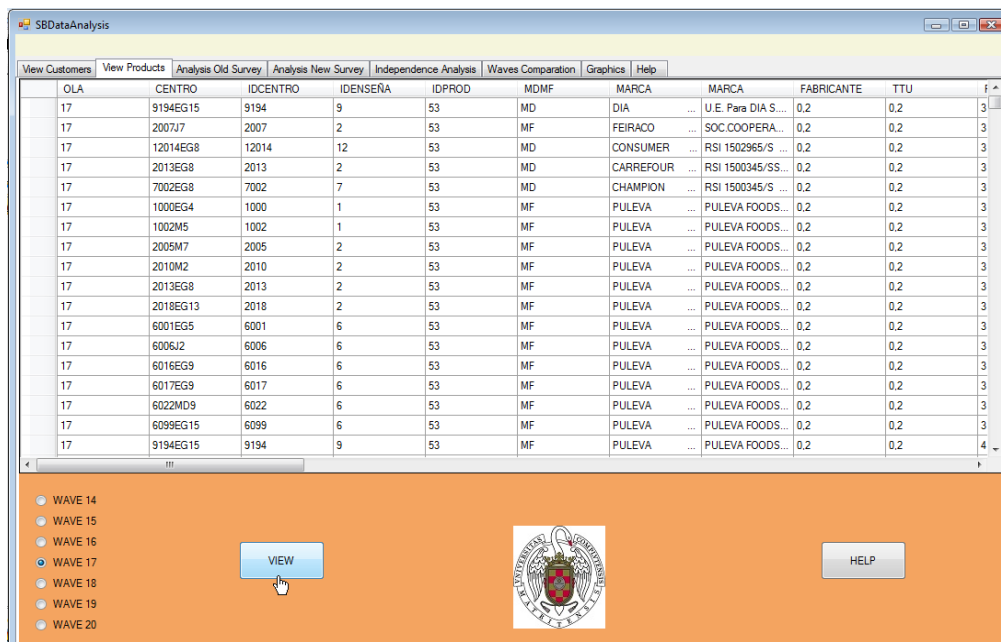


| OLA | CODENCU | IDCENTRO | ENSEÑA | FORMATO | ENCUESTA | DIA        | HORA  | P1 | P2A | P |
|-----|---------|----------|--------|---------|----------|------------|-------|----|-----|---|
| 20  | 1000C   | 1000     | 2      | 1       | C        | 30/11/2007 | 70200 | 1  | 1   | 0 |
| 20  | 1000B   | 1000     | 2      | 1       | B        | 30/11/2007 | 66600 | 1  | 1   | 0 |
| 20  | 1000D   | 1000     | 2      | 1       | D        | 30/11/2007 | 72000 | 1  | 0   | 0 |
| 20  | 1000A   | 1000     | 2      | 1       | A        | 30/11/2007 | 63900 | 1  | 1   | 0 |
| 20  | 1001B   | 1001     | 2      | 1       | B        |            | null  | 1  | 1   | 0 |
| 20  | 1001A   | 1001     | 2      | 1       | A        |            | null  | 1  | 1   | 1 |
| 20  | 1001C   | 1001     | 2      | 1       | C        |            | null  | 1  | 1   | 0 |
| 20  | 1002B   | 1002     | 2      | 1       | B        | 39093      | 47400 | 1  | 1   | 1 |
| 20  | 1002D   | 1002     | 2      | 1       | D        | 39416      | 49380 | 1  | 1   | 1 |
| 20  | 1002A   | 1002     | 2      | 1       | A        | 39093      | 47760 | 1  | 1   | 1 |
| 20  | 1002C   | 1002     | 2      | 1       | C        | 39093      | 48120 | 1  | 1   | 1 |
| 20  | 1003B   | 1003     | 2      | 1       | B        | 26/11/2007 | 61800 | 1  | 1   | 0 |
| 20  | 1003C   | 1003     | 2      | 1       | C        | 26/11/2007 | 62400 | 1  | 1   | 0 |
| 20  | 1003D   | 1003     | 2      | 1       | D        | 26/11/2007 | 63600 | 1  | 1   | 1 |
| 20  | 1003A   | 1003     | 2      | 1       | A        | 26/11/2007 | 61200 | 1  | 0   | 0 |
| 20  | 1003E   | 1003     | 2      | 1       | E        | 26/11/2007 | 64200 | 1  | 1   | 1 |
| 20  | 1004C   | 1004     | 2      | 1       | C        | 29/11/2007 | 70200 | 1  | 1   | 0 |
| 20  | 1004A   | 1004     | 2      | 1       | A        | 29/11/2007 | 68400 | 1  | 1   | 1 |

Figura 15. Se muestra los datos de la encuesta nueva

La segunda pestaña de la visualización, denominada *View Products*, permite al usuario ver los datos de los productos de las distintas olas de las que se dispone. El usuario debe elegir una ola y al pulsar sobre el botón *View* se mostrarán los datos de la B.D de los productos de la ola seleccionada.

En la figura 16 se muestra la visualización de la B.D de productos de las Ola 16.



| OLA | CENTRO   | IDCENTRO | IDENSEÑA | IDPROD | MDMF | MARCA     | MARCA              | FABRICANTE | TTU | F |
|-----|----------|----------|----------|--------|------|-----------|--------------------|------------|-----|---|
| 17  | 9194EG15 | 9194     | 9        | 53     | MD   | DIA       | U.E. Para DIA S... | 0.2        | 0.2 | 3 |
| 17  | 2007J7   | 2007     | 2        | 53     | MF   | FEIRACO   | SOC.COOPERA...     | 0.2        | 0.2 | 3 |
| 17  | 12014EG8 | 12014    | 12       | 53     | MD   | CONSUMER  | RSI 1502965/S...   | 0.2        | 0.2 | 3 |
| 17  | 2013EG8  | 2013     | 2        | 53     | MD   | CARREFOUR | RSI 1500345/S...   | 0.2        | 0.2 | 3 |
| 17  | 7002EG8  | 7002     | 7        | 53     | MD   | CHAMPION  | RSI 1500345/S...   | 0.2        | 0.2 | 3 |
| 17  | 1000EG4  | 1000     | 1        | 53     | MF   | PULEVA    | PULEVA FOODS...    | 0.2        | 0.2 | 3 |
| 17  | 1002M5   | 1002     | 1        | 53     | MF   | PULEVA    | PULEVA FOODS...    | 0.2        | 0.2 | 3 |
| 17  | 2005M7   | 2005     | 2        | 53     | MF   | PULEVA    | PULEVA FOODS...    | 0.2        | 0.2 | 3 |
| 17  | 2010M2   | 2010     | 2        | 53     | MF   | PULEVA    | PULEVA FOODS...    | 0.2        | 0.2 | 3 |
| 17  | 2013EG8  | 2013     | 2        | 53     | MF   | PULEVA    | PULEVA FOODS...    | 0.2        | 0.2 | 3 |
| 17  | 2018EG13 | 2018     | 2        | 53     | MF   | PULEVA    | PULEVA FOODS...    | 0.2        | 0.2 | 3 |
| 17  | 6001EG5  | 6001     | 6        | 53     | MF   | PULEVA    | PULEVA FOODS...    | 0.2        | 0.2 | 3 |
| 17  | 6006J2   | 6006     | 6        | 53     | MF   | PULEVA    | PULEVA FOODS...    | 0.2        | 0.2 | 3 |
| 17  | 6016EG9  | 6016     | 6        | 53     | MF   | PULEVA    | PULEVA FOODS...    | 0.2        | 0.2 | 3 |
| 17  | 6017EG9  | 6017     | 6        | 53     | MF   | PULEVA    | PULEVA FOODS...    | 0.2        | 0.2 | 3 |
| 17  | 6022MD9  | 6022     | 6        | 53     | MF   | PULEVA    | PULEVA FOODS...    | 0.2        | 0.2 | 3 |
| 17  | 6099EG15 | 6099     | 6        | 53     | MF   | PULEVA    | PULEVA FOODS...    | 0.2        | 0.2 | 3 |
| 17  | 9194EG15 | 9194     | 9        | 53     | MF   | PULEVA    | PULEVA FOODS...    | 0.2        | 0.2 | 4 |

Figura 16. Al pulsar sobre el botón *VIEW* se muestran los datos de los productos de la MDD de la Ola escogida

Por último, en ambas pestañas hay un botón Help que permite al usuario ver los significados de las distintas variables de las que esta formado las BB.DD.

### 3.4.2.Módulo 2: Análisis de encuestas de consumidores

En el módulo de análisis de la base de datos, existen dos apartados atendiendo a si se hace referencia a las encuestas antiguas (OLAS 14-19), o a las encuestas nuevas (a partir OLA 20), ya que los datos variaban de uno a otro tipo de encuestas. Esto se traduce en que se ha tenido que diferenciar los estudios, teniendo que hacer dos pestañas diferentes pero que tendrán el mismo fin. A continuación se hará una explicación conjunta de ambas, señalando las similitudes pero mostrando imágenes de ambas pestañas con ejemplos de la ejecución.

A continuación se puede observar el aspecto de la pestaña para el análisis de las OLAS antiguas. Se puede ver el conocido botón *Help* que guiará si se tiene duda sobre cómo usar la pestaña, y también se puede observar en el centro de la ventana un desplegable que tiene diferentes casos de estudio. En la figura 17 se puede observar los diferentes estudios disponibles para las encuestas antiguas a consumidores

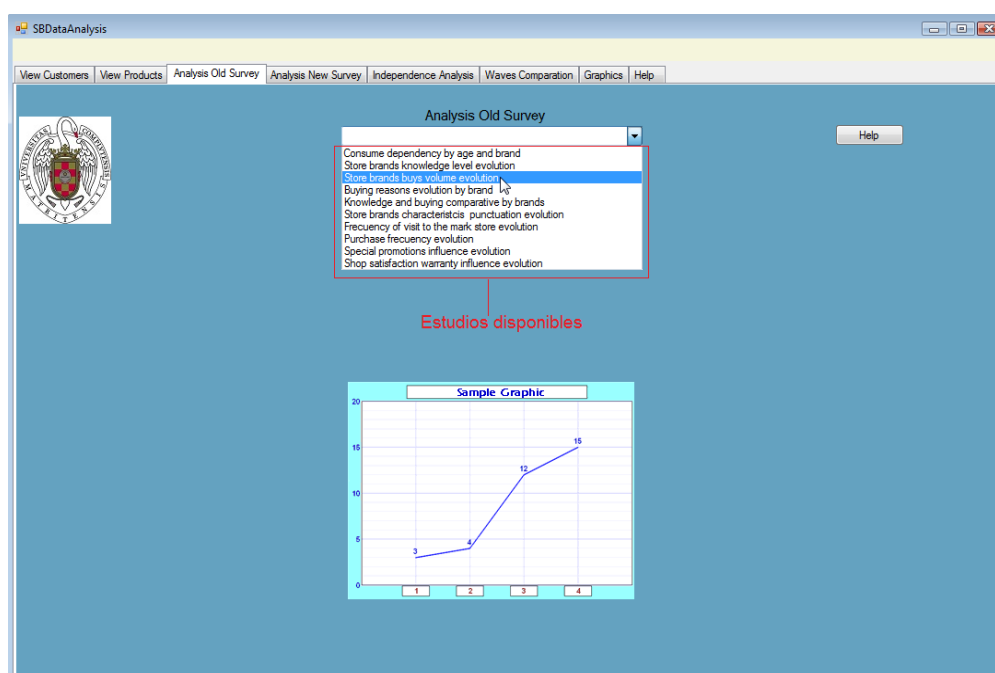


Figura 17. Pestaña análisis encuesta antigua

No todos los estudios se realizan sobre las mismas variables, por lo que en función del estudio se tendrá unas opciones u otras para el estudio. Dichas opciones serán las variables en función de las que realizará el estudio, a parte de las que ya vienen implícitas en la selección del estudio, como por ejemplo se hará el estudio *%Store brands buys volume evolution+*, y que incluirá las variables OLA y compra ya que sobre eso trata el estudio. Pero dentro del estudio puede ser interesante centrarse en uno u otro aspecto, y por eso aparecen las diferentes opciones (en este caso enseña, edad, género y estatus social), como se muestra en la figura 18, para centrarse en una característica concreta de los compradores.

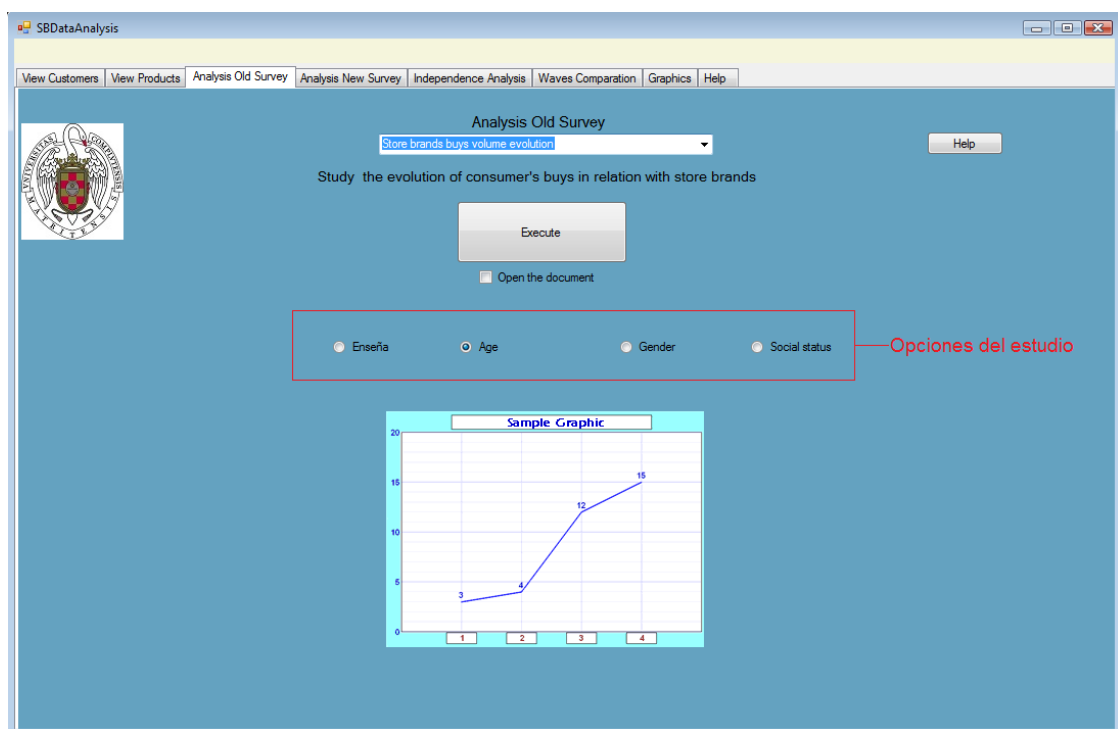
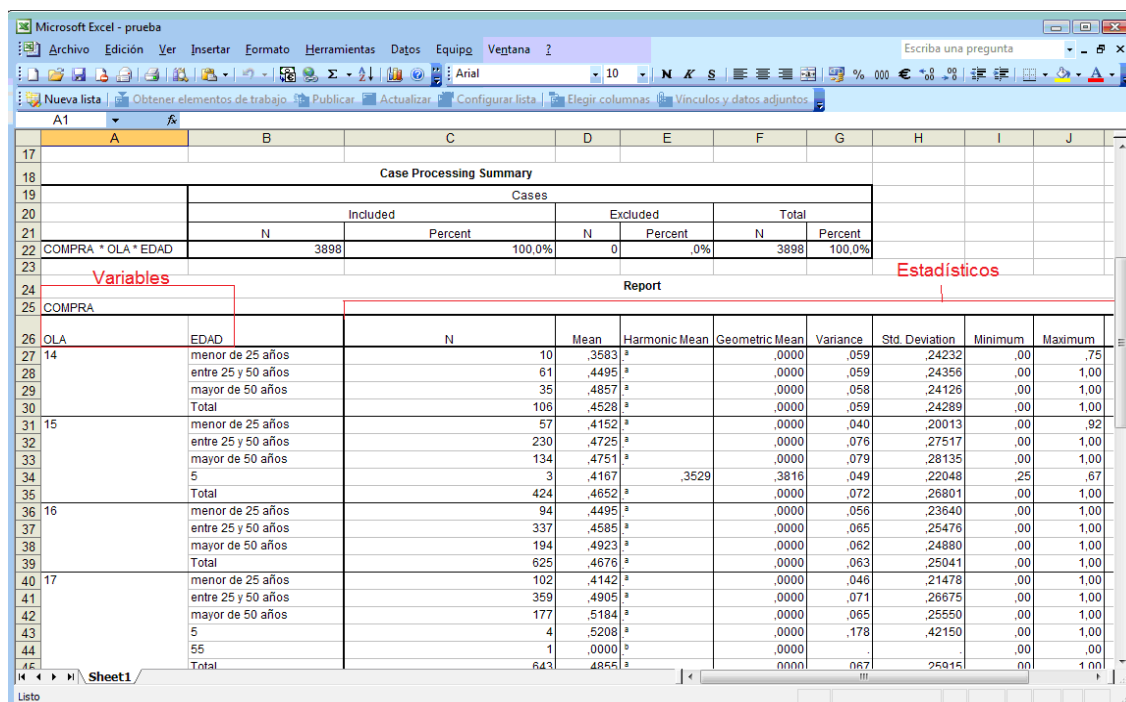


Figura 18. Diferentes opciones que hay para el estudio

El estudio realizado se guardará en un archivo de tipo Excel, dónde el usuario indique y con el nombre que desee. Una vez realizado el estudio se en el caso de que el usuario hay seleccionado *Open the document* se mostrará el archivo con los resultados del estudio, como se muestra en la figura 19.



| Case Processing Summary |      |          |   |          |      |         |  |  |  |  |
|-------------------------|------|----------|---|----------|------|---------|--|--|--|--|
|                         |      | Included |   | Excluded |      | Total   |  |  |  |  |
|                         | N    | Percent  | N | Percent  | N    | Percent |  |  |  |  |
| COMPRA * OLA * EDAD     | 3898 | 100,0%   | 0 | ,0%      | 3898 | 100,0%  |  |  |  |  |

| Report |                    |     |                    |               |                |          |                |         |         |  |
|--------|--------------------|-----|--------------------|---------------|----------------|----------|----------------|---------|---------|--|
| OLA    | EDAD               | N   | Mean               | Harmonic Mean | Geometric Mean | Variance | Std. Deviation | Minimum | Maximum |  |
| 14     | menor de 25 años   | 10  | ,3583 <sup>a</sup> |               | ,0000          | ,059     | ,24232         | ,00     | ,75     |  |
|        | entre 25 y 50 años | 61  | ,4495 <sup>a</sup> |               | ,0000          | ,059     | ,24356         | ,00     | 1,00    |  |
|        | mayor de 50 años   | 35  | ,4857 <sup>a</sup> |               | ,0000          | ,058     | ,24126         | ,00     | 1,00    |  |
|        | Total              | 106 | ,4528 <sup>a</sup> |               | ,0000          | ,059     | ,24289         | ,00     | 1,00    |  |
|        |                    | 57  | ,4152 <sup>a</sup> |               | ,0000          | ,040     | ,20013         | ,00     | ,92     |  |
| 15     | menor de 25 años   | 230 | ,4725 <sup>a</sup> |               | ,0000          | ,076     | ,27517         | ,00     | 1,00    |  |
|        | entre 25 y 50 años | 134 | ,4751 <sup>a</sup> |               | ,0000          | ,079     | ,28135         | ,00     | 1,00    |  |
|        | mayor de 50 años   | 3   | ,4167 <sup>a</sup> | ,3529         | ,3816          | ,049     | ,22048         | ,25     | ,67     |  |
|        | Total              | 424 | ,4652 <sup>a</sup> |               | ,0000          | ,072     | ,26801         | ,00     | 1,00    |  |
|        |                    | 94  | ,4495 <sup>a</sup> |               | ,0000          | ,056     | ,23640         | ,00     | 1,00    |  |
| 16     | menor de 25 años   | 337 | ,4585 <sup>a</sup> |               | ,0000          | ,065     | ,25476         | ,00     | 1,00    |  |
|        | entre 25 y 50 años | 194 | ,4923 <sup>a</sup> |               | ,0000          | ,062     | ,24880         | ,00     | 1,00    |  |
|        | mayor de 50 años   | 625 | ,4676 <sup>a</sup> |               | ,0000          | ,063     | ,25041         | ,00     | 1,00    |  |
|        | Total              | 102 | ,4142 <sup>a</sup> |               | ,0000          | ,046     | ,21478         | ,00     | 1,00    |  |
|        |                    | 359 | ,4905 <sup>a</sup> |               | ,0000          | ,071     | ,26675         | ,00     | 1,00    |  |
| 17     | menor de 25 años   | 177 | ,5184 <sup>a</sup> |               | ,0000          | ,065     | ,25550         | ,00     | 1,00    |  |
|        | entre 25 y 50 años | 4   | ,5208 <sup>a</sup> |               | ,0000          | ,178     | ,42150         | ,00     | 1,00    |  |
|        | mayor de 50 años   | 1   | ,0000 <sup>a</sup> |               | ,0000          |          |                | ,00     | ,00     |  |
|        | 55                 | 643 | ,4855 <sup>a</sup> |               | ,0000          | ,067     | ,25915         | ,00     | 1,00    |  |
|        | Total              |     |                    |               |                |          |                |         |         |  |

Figura19. Resultados del estudio almacenado en un Excel

En dicho archivo se mostrarán, atendiendo al criterio que se ha elegido en el apartado de opciones, una tabla con los resultados ordenados. Como se puede observar, el archivo muestra la tabla de un estudio estadístico completo primero ordenado por OLA y compra, y como parámetro final la edad. Dicho estudio, mostrará el número de elementos que se estudian ordenados por los criterios anteriores, también se puede ver los estadísticos típicos como la media, la varianza, la desviación, el máximo y el mínimo, y en caso de que fuese posible también se mostrará las medias armónica y geométrica para poder sacar datos concluyentes.

Con todo esto el estudio podría darse por concluido, pero para una mejor visualización, y para ver los datos obtenidos más directamente, se generan unos gráficos (de 1 a 6 dependiendo del estudio que se realice), donde se mostrará de forma visual los resultados obtenidos en la tabla del documento Excel, como se muestra en la figura 20. Dichos gráficos se encuentran en la pestaña, y en caso de que resulte demasiado pequeño el gráfico en la ventana, por ejemplo en la situación en que se muestran 6 gráficos, se puede pulsar sobre uno de ellos para que se muestre en grande en toda la ventana en la pestaña *Graphics*.

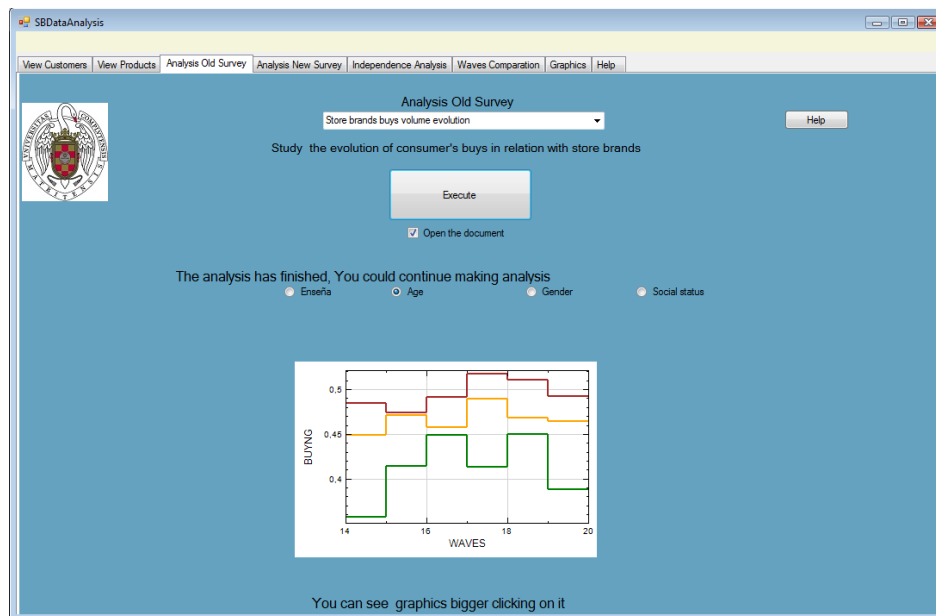


Figura 20. Al acabar el estudio se muestran gráficos con los resultados

Como se ha comentado anteriormente se dispone de una segunda pestaña para realizar estudios de las encuestas nuevas. Dicha pestaña se encuentra a continuación de la pestaña de análisis de la encuesta antigua, y como se puede observar, se realizan estudios diferentes. También se puede observar que las opciones para el estudio son diferentes, y esto es debido a que las encuestas no tienen exactamente los mismos campos, por lo que da lugar a estudios diferentes.

En la figura 21 se muestra los distintos estudios disponibles para las encuestas nuevas.

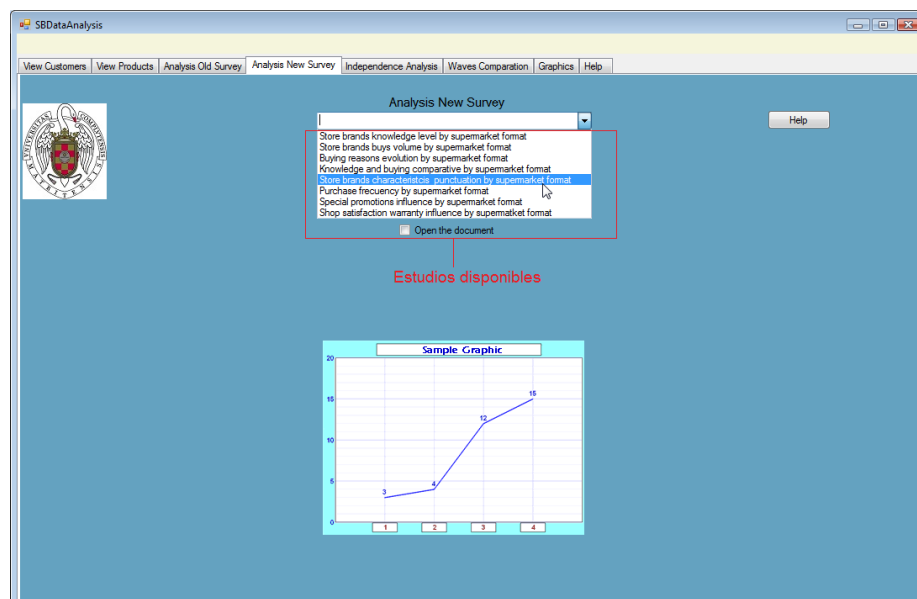


Figura 21. Estudios disponibles en la pestaña de análisis de encuesta nueva

En la figura 22 se muestra las distintas opciones que tiene el estudio.

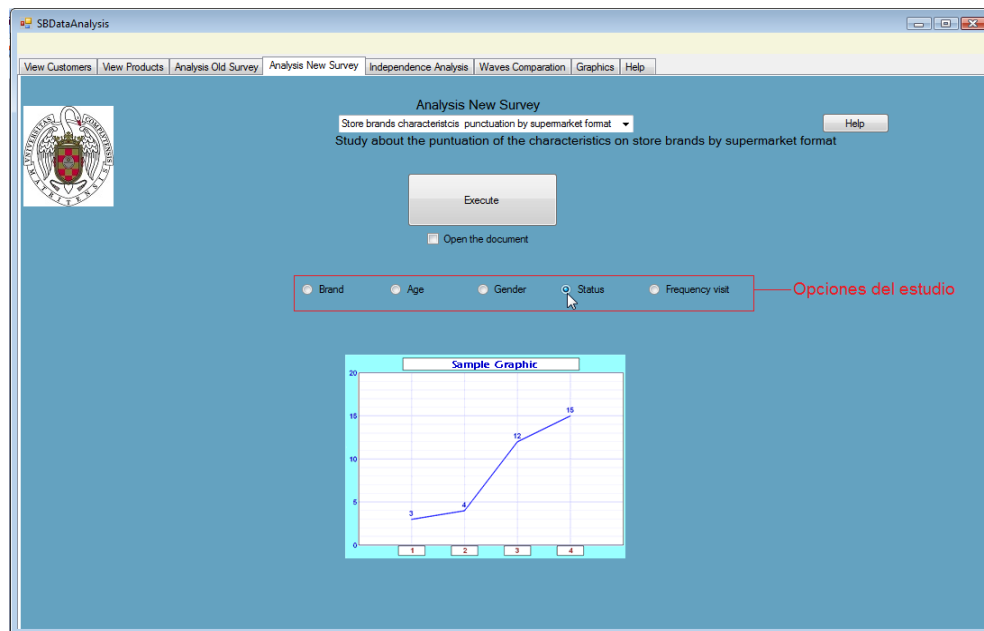


Figura 22. Opciones disponibles para cada estudio

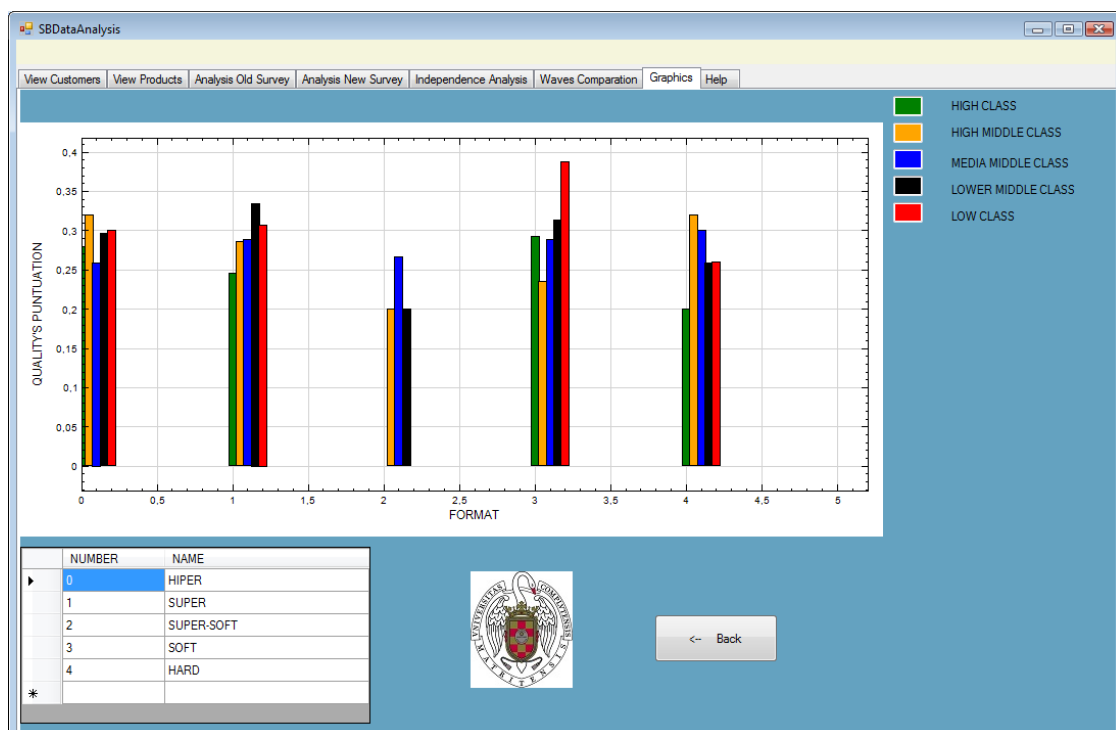
En el estudio de la encuesta nueva que se muestra en la figura 23, se observa que salen varias gráficas, ya que el campo %Formato+ tiene 5 posibles selecciones en el cuestionario, y habrá que realizar el estudio para cada una de ellas. Esto viene bien debido a que así se puede mostrar mejor la visión de la ventana cuando se selecciona una gráfica para mostrarla en toda la ventana.



Figura 23. Al acabar el estudio se muestran gráficos con los resultados

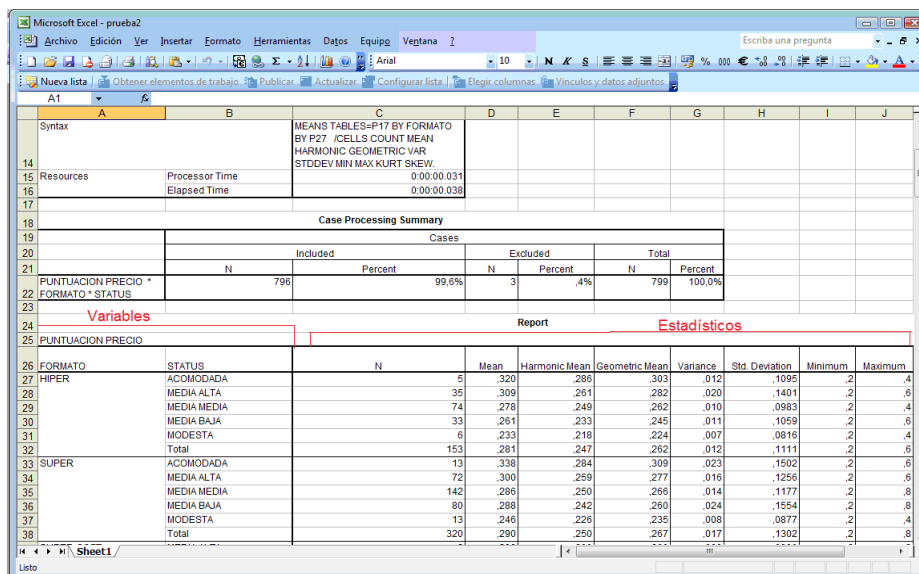
Cuando se accede a la vista del gráfico se muestra no solo la gráfica que se ha seleccionado, si no que se incluirá una leyenda (en el margen superior derecho), que indicará el significado de cada color, como se muestra en la figura 24.

Además, en este estudio, se ha usado la variable %Status+, que tiene 5 posibles valores diferentes no excluyentes, y que serán los que se indiquen en una segunda leyenda situada en la parte inferior izquierda de la ventana, y que como se ve hará referencia a cada uno de los grupos de barras, que de otra forma no se sabría lo que significan, pero que de este modo se puede diferenciar.



**Figura 24.** Gráfico ampliado con leyenda y significado de los valores de las abcisas

Pulsando sobre el botón BACK se volverá a la pestaña de dónde es el gráfico ampliado, desde donde podremos ver de nuevo todas las gráficas y seleccionar otra. Como se ve para el apartado de las encuestas antiguas, también se mostrará el documento excel del estudio con la tabla ordenada por variables, como se muestra en la figura 25.



MEANS TABLES=P17 BY FORMATO  
BY P27 / CELLS COUNT MEAN  
HARMONIC GEOMETRIC VAR  
STDDEV MIN MAX KURT SKEW.

Resources Processor Time 0:00:00.031  
Elapsed Time 0:00:00.038

Case Processing Summary

|   | Included |         | Excluded |         | Total |         |
|---|----------|---------|----------|---------|-------|---------|
|   | N        | Percent | N        | Percent | N     | Percent |
| PUNTUACION PRECIO *<br>FORMATO * STATUS | 796      | 99.6%   | 3        | .4%     | 799   | 100.0%  |

Variables Report Estadísticos

|         |             | N   | Mean | Harmonic Mean | Geometric Mean | Variance | Std. Deviation | Minimum | Maximum |
|---------|-------------|-----|------|---------------|----------------|----------|----------------|---------|---------|
| FORMATO | STATUS      |     |      |               |                |          |                |         |         |
| HIPER   | ACOMODADA   | 5   | .320 | .286          | .303           | .012     | .1095          | .2      | .4      |
|         | MEDIA ALTA  | 35  | .309 | .261          | .282           | .020     | .1401          | .2      | .6      |
|         | MEDIA MEDIA | 74  | .278 | .249          | .262           | .010     | .0983          | .2      | .4      |
|         | MEDIA BAJA  | 33  | .261 | .233          | .245           | .011     | .1059          | .2      | .6      |
|         | MODESTA     | 6   | .233 | .219          | .224           | .007     | .0616          | .2      | .4      |
|         | Total       | 153 | .281 | .247          | .262           | .012     | .1111          | .2      | .6      |
| SUPER   | ACOMODADA   | 13  | .338 | .284          | .309           | .023     | .1502          | .2      | .6      |
|         | MEDIA ALTA  | 72  | .300 | .259          | .277           | .016     | .1256          | .2      | .6      |
|         | MEDIA MEDIA | 142 | .286 | .250          | .266           | .014     | .1177          | .2      | .8      |
|         | MEDIA BAJA  | 80  | .288 | .242          | .260           | .024     | .1554          | .2      | .8      |
|         | MODESTA     | 13  | .246 | .226          | .235           | .008     | .0877          | .2      | .4      |
|         | Total       | 320 | .290 | .250          | .267           | .017     | .1302          | .2      | .8      |

Figura25. Resultados del estudio almacenado en un Excel

### 3.4.3. Módulo 3: Análisis de Independencia

Uno de los puntos de interés más importantes para el proyecto y para cualquiera que esté interesado trabajar con la base de datos proporcionada es conocer que variables que resultan determinantes a la hora de adquirir o no un producto están relacionadas entre sí.

Por ejemplo, ¿Es cierto que la edad y el conocimiento sobre un determinado producto están relacionados?, ¿Se puede determinar que según el género de la persona la distancia al supermercado es un factor importante?, ¿El status social influye a la hora de dejarse llevar por el precio?

En el apartado de análisis de independencia se ofrece una respuesta a estas cuestiones y a muchas más. Tal y como se muestra en la figura 26, al trabajar con dos bases de datos distintas, una para las olas de la 14 a la 19 y otra para las olas siguientes, se pueden realizar 240 estudios distintos para las olas antiguas y 272 para la nueva ola. Esto da un total de 512 posibles comparaciones para obtener conclusiones.



Otra ventaja a su vez de contar con bases de datos diferenciadas es que se puede comprobar si con el paso de los años 2 variables aparentemente independientes entre sí, arrojan un resultado contrario y se convierten en dependientes. Esto es posible por muchos motivos, cambios en la sociedad, mejora de la toma de datos, etc.

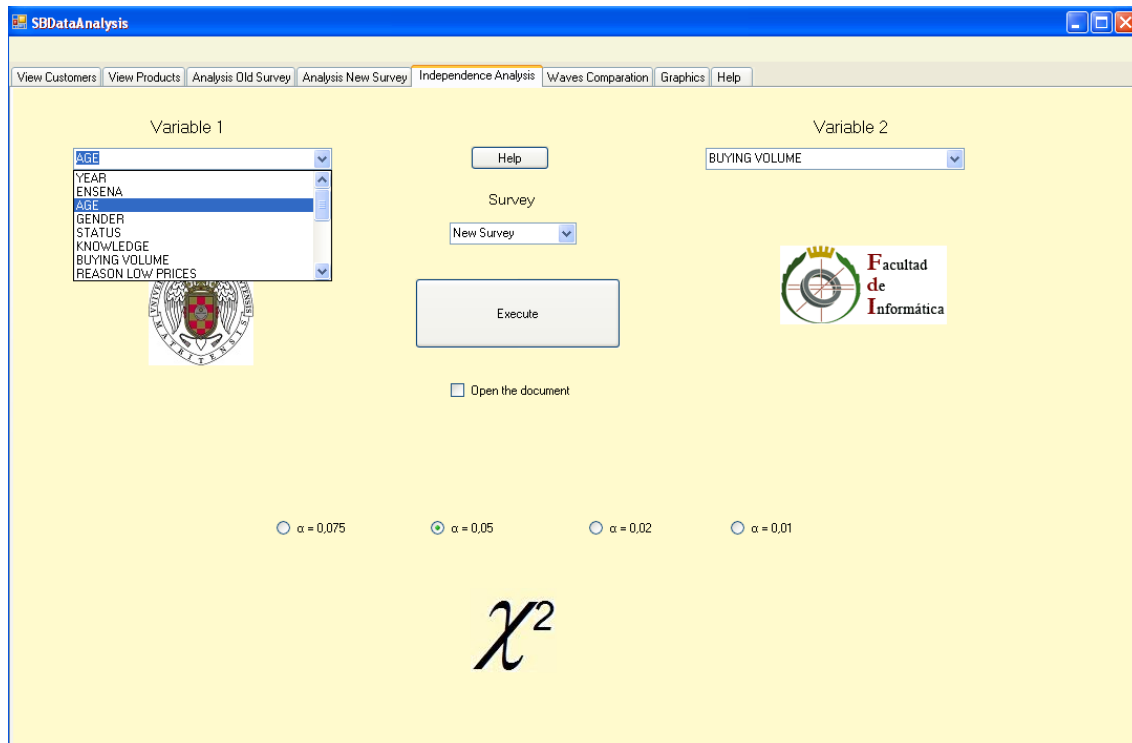


Figura 26. Pestaña de análisis de independencia.

Como se puede comprobar en la imagen superior, la figura 25, para realizar los estudios existen dos listas desplegables, variable 1 y variable 2, en ellas se pueden escoger entre las opciones de variables disponibles en la base de datos para realizar el estudio.

Según se muestra en la figura 27, en medio de los dos desplegables de variables, existe otro desplegable más pequeño, el desplegable de los estudios, o survey. En él se podrá elegir entre los estudios antiguos (olas de la 14 a la 19) y el nuevo estudio, ola 20. La diferencia entre los estudios antiguos y el nuevo, es la variación de formato de las encuestas, así como que el nuevo estudio posee una variable nueva llama *formato*, que define el tipo de supermercado en la que se realiza la encuesta: Hipermercado, Supermercado, Súper-soft, soft o Hard.

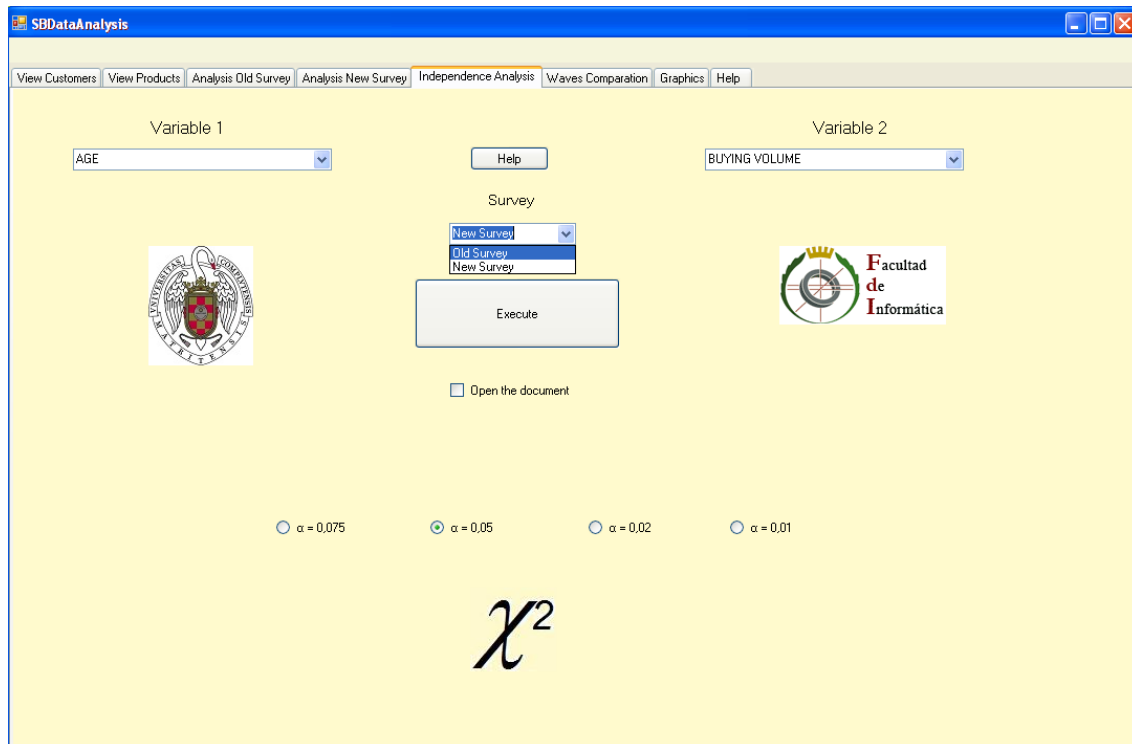


Figura 27. Pestaña de análisis de independencia.

Para determinar si dos variables son independientes entre sí o no se realiza estudios de chi-cuadrado (chi-square). Cuya fórmula se define así y que se ha explicado anteriormente en el apartado de conceptos.

$$\chi^2 \approx \sum_{ij} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Se acepta la hipótesis nula si  $\chi^2 < \chi^2_{1-\alpha, v}$ , el percentil 1 -  $\alpha$  de la distribución chi-cuadrado con  $v$  grados de libertad.

Para realizar estos estudios se permite dar varios posibles valores de alpha, de forma que esta puede ser más o menos restrictiva según las necesidades del usuario.

Una vez fijados todos los parámetros, lo único que se debe hacer es apretar el botón de ejecutar, para generar los resultados. Se abrirá una pestaña para indicar en qué lugar se desea guardar el archivo Excel que se generará. Si tenemos activa la casilla de *Open the document*, cuando acabe la ejecución el fichero generado se abrirá automáticamente y se obtendrá un resultado como que se muestra a continuación.

|    | A  | B   | C                               | D     |
|----|--|---|---------------------------------|-------|
| 4  | <b>Notes</b>   |   |                                 |       |
| 5  | Output Created   | 15-may-2010 10:20:07  |                                 |       |
| 6  | Comments   |   |                                 |       |
| 7  | Input  | Data  | C:\INFORMÁTICA\compradoresOLA20 |       |
| 8  |  | Filter  | <none>                          |       |
| 9  |  | Weight  | <none>                          |       |
| 10 |  | Split File  | <none>                          |       |
| 11 |  | N of Rows in Working Data File  | 799                             |       |
| 12 | Syntax   | CTABLES /VARIABLES<br>VARIABLES=P27 P14B<br>DISPLAY=LABEL /TABLE P27<br>[C][COUNT F40.0] BY P14B [C]<br>/CATEGORIES VARIABLES=P27 P14B<br>ORDER=A KEY=VALUE<br>EMPTY=EXCLUDE /SIGTEST<br>TYPE=CHISQUARE ALPHA=0.05<br>INCLUDEMRSETS=YES |                                 |       |
| 13 | Resources  | Processor Time  | 0:00:00.015                     |       |
| 14 |  | Elapsed Time  | 0:00:00.022                     |       |
| 15 |  |   |                                 |       |
| 16 |  |   |                                 |       |
| 17 | Table 1  |   |                                 |       |
| 18 |  |   |                                 |       |
| 19 |  |   | <b>RAZON DISTANCIA</b>          |       |
| 20 |  |   | 0                               | 1     |
| 21 |  |   | Count                           | Count |
| 22 | STATUS   | ACOMODADA   | 9                               | 23    |
| 23 |  | MEDIA ALTA  | 47                              | 126   |
| 24 |  | MEDIA MEDIA   | 89                              | 280   |
| 25 |  | MEDIA BAJA  | 35                              | 142   |
| 26 |  | MODESTA   | 11                              | 34    |
| 27 |  |   |                                 |       |
| 28 | <b>Pearson Chi-Square Tests</b>  |   |                                 |       |
| 29 |  |   | <b>RAZON DISTANCIA</b>          |       |
| 30 | STATUS   | Chi-square  | 2,991                           |       |
| 31 |  | df  | 4                               |       |
| 32 |  | Sig.  | ,559                            |       |
| 33 | Results are based on nonempty rows and columns in each innermost subtable. |   |                                 |       |

**Figura28.** Resultados análisis independencia entre Status social y distancia en los últimos sondeos

Tal y como se muestra en la figura 28, la imagen corresponde a la captura de un estudio de independencia entre el status social y la razón de distancia.

La primera tabla contiene información meramente técnica, como la base de datos que se ha usado (la de la ola20), filas de la base de datos, filtros empleados, tiempo empleado en la ejecución del estudio, etc.

La segunda tabla muestra los datos recogidos en sí, separados por status social, recoge las respuestas almacenadas en la base de datos.

La tercera tabla muestra los resultados obtenidos, de los tres que se muestran, el más interesante para el estudio es el parámetro Sig. Que es el correspondiente al Alpha, si  $\text{Sig} < \text{Alpha}$ , en este caso 0,05 serán independientes. En nuestro caso al ser  $\text{Sig} > \text{Alpha}$ , se puede determinar que las variables son dependientes Sin embargo, la figura 29 muestra que este mismo estudio para olas anteriores, y se comprueba que los resultados obtenidos son distintos. Una muestra más de lo volátil que puede ser nuestra sociedad.

|    | A  | B                              | C  | D     | E |
|----|--|--------------------------------|--|-------|---|
| 4  | Notes  |                                |  |       |   |
| 5  | Output Created   |                                | 15-may-2010 10:20:34   |       |   |
| 6  | Comments   |                                |  |       |   |
| 7  | Input  | Data                           | C:\INFORMÁTICA\compradores.sav   |       |   |
| 8  |  | Filter                         | <none>   |       |   |
| 9  |  | Weight                         | <none>   |       |   |
| 10 |  | Split File                     | <none>   |       |   |
| 11 |  | N of Rows in Working Data File | 3898   |       |   |
| 12 | Syntax   |                                | CTABLES /MLABELS<br>VARIABLES=P27 P14B<br>DISPLAY=LABEL /TABLE P27<br>[C]COUNT F40.0] BY P14B [C]<br>/CATEGORIES VARIABLES=P27 P14B<br>ORDER=A KEY=VALUE<br>EMPTY=EXCLUDE /SIGTEST<br>TYPE=CHISQUARE ALPHA=0.05<br>INCLUDEMRSETS=YES |       |   |
| 13 | Resources  | Processor Time                 | 0:00:00.031  |       |   |
| 14 |  | Elapsed Time                   | 0:00:00.034  |       |   |
| 15 |  |                                |  |       |   |
| 16 |  |                                |  |       |   |
| 17 | Table 1  |                                |  |       |   |
| 18 |  |                                |  |       |   |
| 19 |  |                                | RAZÓN DISTANCIA  |       |   |
| 20 |  |                                | 0  | 1     |   |
| 21 |  |                                | Count  | Count |   |
| 22 | STATUS   | ACOMODADA                      | 55   | 125   |   |
| 23 |  | MEDIA-ALTA                     | 265  | 751   |   |
| 24 |  | MEDIA-MEDIA                    | 426  | 1413  |   |
| 25 |  | MEDIA-BAJA                     | 132  | 529   |   |
| 26 |  | MODESTA                        | 37   | 108   |   |
| 27 |  |                                |  |       |   |
| 28 | Pearson Chi-Square Tests   |                                |  |       |   |
| 29 |  |                                | RAZÓN DISTANCIA  |       |   |
| 30 | STATUS   | Chi-square                     | 13,431   |       |   |
| 31 |  | df                             | 4  |       |   |
| 32 |  | Sig.                           | ,009   |       |   |
| 33 | Results are based on nonempty rows and columns in each innermost subtable. |                                |  |       |   |
|    | *. The Chi-square statistic is significant at the 0.05 level.              |                                |  |       |   |

**Figura29.** Comparación entre el status social y el distancia en los primeros estudios.

Se puede comprobar que en este estudio, realizado para los mismos parámetros pero con otra base de datos, se obtiene que son independientes. Es por ello muy importante contar con buenas cantidades de datos recogidos, ya que se necesita una gran precisión para detectar a tiempo los cambios por los que pasa la sociedad.

Por último, en la pestaña se puede observar a la derecha un botón help, el cual dará una descripción detallada del uso de dicha pestaña.

### 3.4.4. Modulo 4: Comparación de consumo usando métodos de decisión multicriterio

En este apartado del proyecto el objetivo del estudio es determinar el sector generacional más susceptible de compra sobre el cual enfocar futuras campañas publicitarias para las MDD. Los resultados obtenidos se basan en los datos almacenados en OLAS (14 a 19). En la figura 30 se puede observar el aspecto de la pestaña que realiza el estudio, a la que hará referencia durante su explicación.

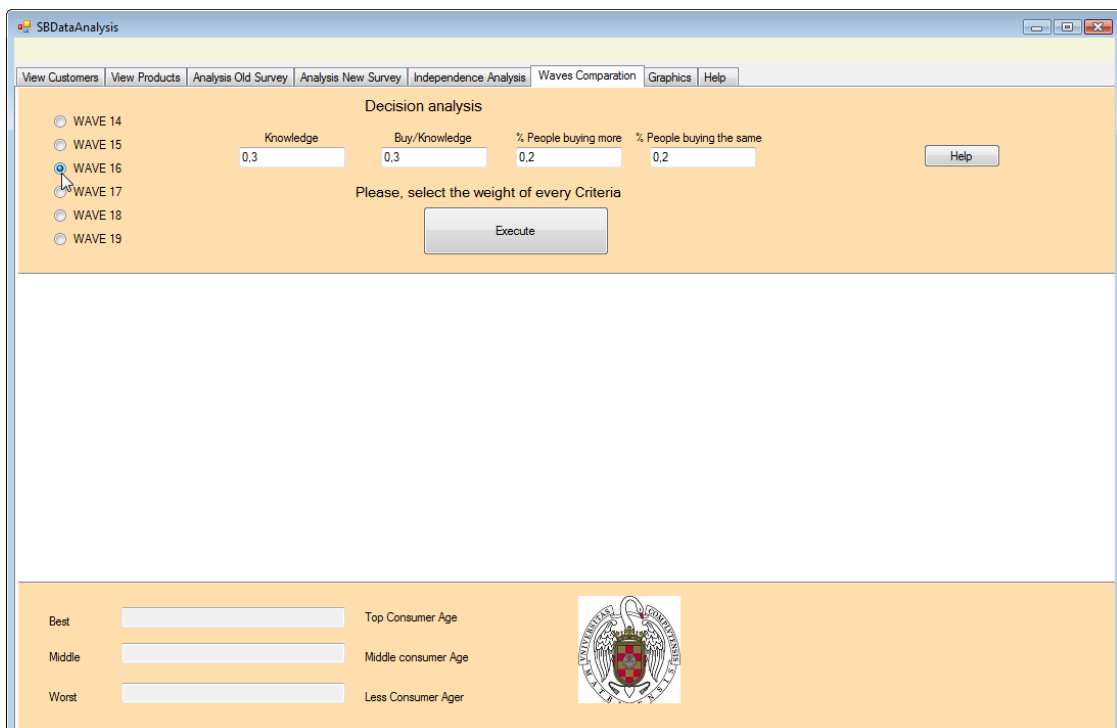


Figura 30. Pestaña de comparación de consumo

Durante el desarrollo de la herramienta se han utilizado diversos sistemas clásicos de decisión relativos a la inferencia estadística como el test de la Chi-cuadrado entre otros. Sin embargo la gran cantidad de variables cualitativas de interés ha obligado a incluir también sistemas de decisión basados en la incertidumbre. En este trabajo se muestra en concreto la solución que se ha decidido lleva a cabo para el análisis del comportamiento del consumidor utilizando lógica difusa y técnicas de Soft Computing.

La elaboración de un patrón de comportamiento sobre hábitos de consumo de los sectores poblacionales de acuerdo a tres rangos de edad ha sido uno de los primeros pasos a realizar. Al ser datos referidos al consumo de MDD en la habitual cesta de la compra, se han clasificado los siguientes grandes grupos:

1. Menores de 25 años
2. Entre 25 y 50 años
3. Mayores de 50 años

Debido a que las costumbres pueden variar mucho entre generaciones y más teniendo en cuenta el espectacular auge de nuevas tecnologías y formas de comunicación que son mucho más fáciles de orientar a unos rangos de edad que a otros, el posicionamiento de las marcas comerciales se basa en la realización de campañas publicitarias dirigidas a sectores concretos de población. De esta forma las compañías pueden no sólo aumentar el número de ventas sino también realizar mantenimiento de consumidores y adaptación de productos a las nuevas tendencias.

La base de datos original posee un alto grado de incertidumbre debido a varios factores, entre ellos, un gran número de variables cualitativas. El uso de técnicas de Soft Computing para el análisis de esta BBDD es justificable. La herramienta SBDataAnalysis utiliza un método de decisión multicriterio (MCDM) basado en puntos óptimos. Este método se explica con detalle en la sección de conceptos teóricos aplicados.

En la pestaña se puede observar a la derecha un botón help, el cual dará una descripción detallada del uso de dicha pestaña. En la parte izquierda de la pantalla se encuentran los botones para seleccionar la OLA sobre la que se quiere realizar el estudio. Debido a las características del estudio solo se puede seleccionar una OLA simultáneamente, por lo que al seleccionar una se quitará la selección de alguna otra que se hubiera elegido anteriormente.

En el centro de la ventana hay 4 ventanas, las cuales servirán para indicar los pesos de las diferentes variables. Estos valores de dichos campos deben tener una suma de 1 ya que así será correcta la transformación mediante los pesos. Estos valores se tratan como variables ya que de esta forma en un futuro se podrá variar la importancia de cada variable atendiendo a lo que mejor se piense que se adapta para el estudio en el futuro.

Tras ajustar todos los valores de los pesos correctamente, y seleccionar la OLA que deseamos estudiar, se debe pulsar sobre el botón %Execute+, el cual dará paso a la iniciación del programa. Tras esto, se puede ver que el panel que ocupa la zona central de la ventana se ha rellenado con los valores de las tablas que se describió en el apartado de teoría sobre el método de decisión multicriterio, y que representan todos los valores intermedios que se debe tomar para la obtención del resultado. Los resultados obtenidos en la ventana no son exactamente los de las tablas que se muestran antes, debido a que el programa SPSS realiza un redondeo más exacto que el que se ha podido conseguir estudiando los datos manualmente, pero viendo que los resultados son muy similares se puede observar que el procedimiento se ejecuta correctamente.

El resultado obtenido del estudio de las tablas, como se muestra en la figura 31, se representará en los tres apartados que se encuentran en la parte inferior, los cuales mostrarán los datos  $p_i$  ordenados de modo que se podrá ver que decisión es la mejor. Para no incurrir en confusión está descrito al lado cual es la mejor opción, atendiendo a la pregunta (ya que si se cambiara la pregunta, los resultados se nos representarían invertidos y podría dar pie a confusión)

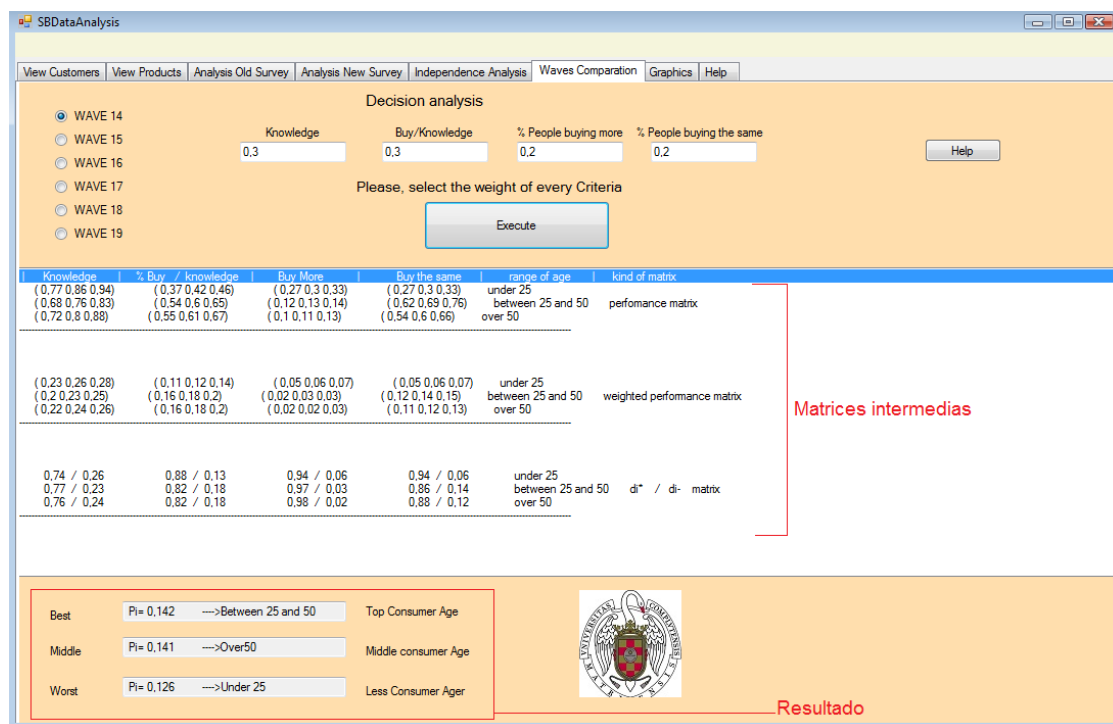
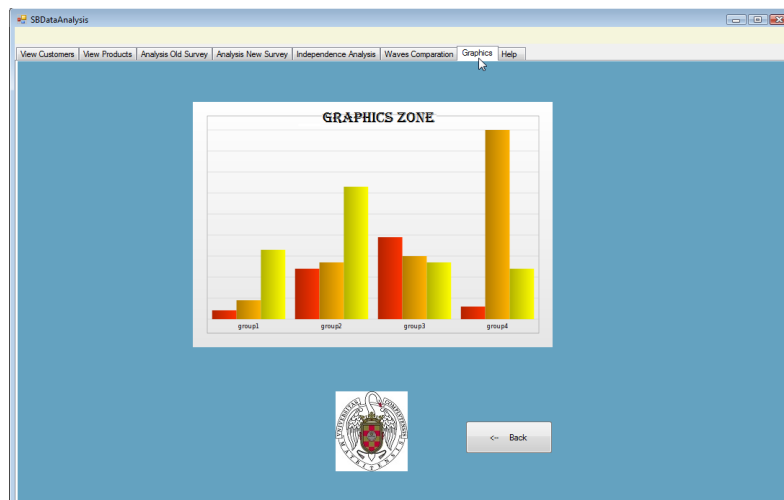


Figura31. Comparación de consumo de la Ola14

### 3.4.5. Módulo 5: Gráficos

En la pestaña *Graphics* es donde se muestran los gráficos que el usuario quiere ver en más detalle, en ella se muestra el gráfico seleccionado más grande, su leyenda y, si procede, una explicación de los valores de las abscisas (en el caso de enseñas y formatos). Los gráficos mostrados provienen de los diferentes estudios de análisis de encuestas, tanto antiguas como nuevas.

En la figura 32 se observa la pestaña Graphics antes de seleccionar un gráfico para ser visto ampliado.



**Figura32.** La ventana de gráficos antes de hacer cualquier estudio de encuestas

Una vez que el usuario ha pulsado sobre un gráfico, creado al realizar un estudio, se muestra la pestaña Graphics, en la que se muestra ese gráfico con más detalle. En la figura 33 se observa cómo hacer para ver un gráfico ampliado, simplemente pulsando sobre el gráfico.



**Figura 33.** Para ver un gráfico con más detalle se debe pulsar sobre él

Una vez pulsado sobre el gráfico, la aplicación se dirige a la pestaña Graphics y muestra el gráfico seleccionado ampliado, mostrando también su leyenda y, si procede, el significado de los valores de las abscisas; como se muestra en la figura 34.



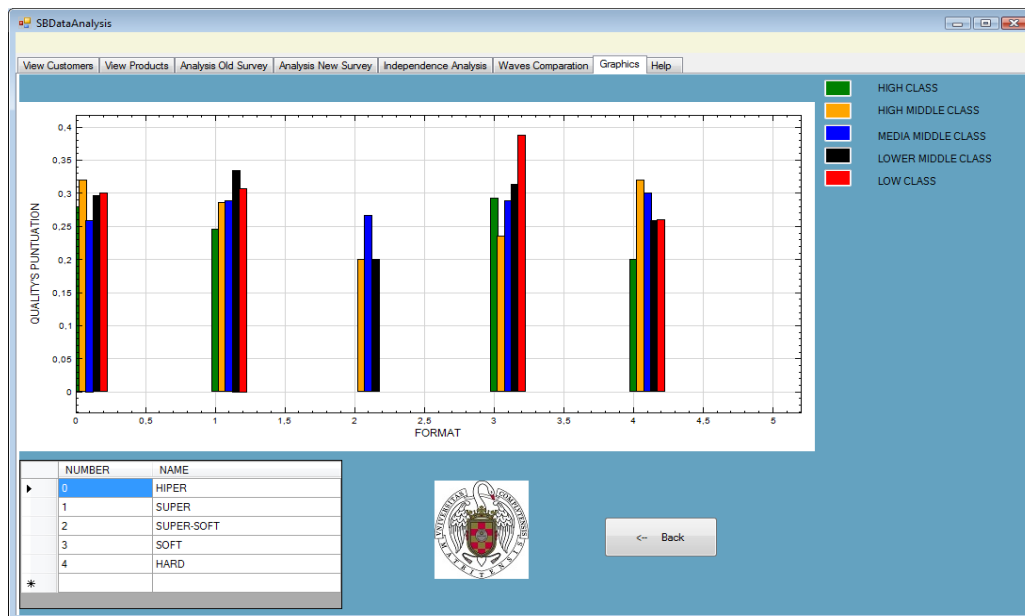


Figura 34. La pestaña Graphics al seleccionar un gráfico, mostrando la leyenda y el significado de las abscisas

Como se puede observar en la pestaña hay un botón back que nos devuelve a la pestaña desde la cuál se ha pulsado sobre el gráfico que se quería ver en detalle.

### 3.4.6. Módulo 6: Ayuda

Todas las pestañas del proyecto, menos la que muestra los gráficos ampliados cuentan con un botón llamado *Help*, la misión de este botón es abrir la pestaña de nombre homónimo. En ella dependiendo del lugar en el que haya sido llamado se mostrará un texto en inglés, como el resto del proyecto, con instrucciones precisas de cómo utilizar la pestaña en cuestión o mostrar el significado de las variables en el caso de las pestañas de visualización.

A continuación se muestra un ejemplo, el programa se encuentra en la pestaña de análisis de datos de la encuesta nueva de consumidores (new survey), y se quiere saber cómo utilizarla correctamente, se debe pulsar sobre el botón help como se muestra en la figura 35.

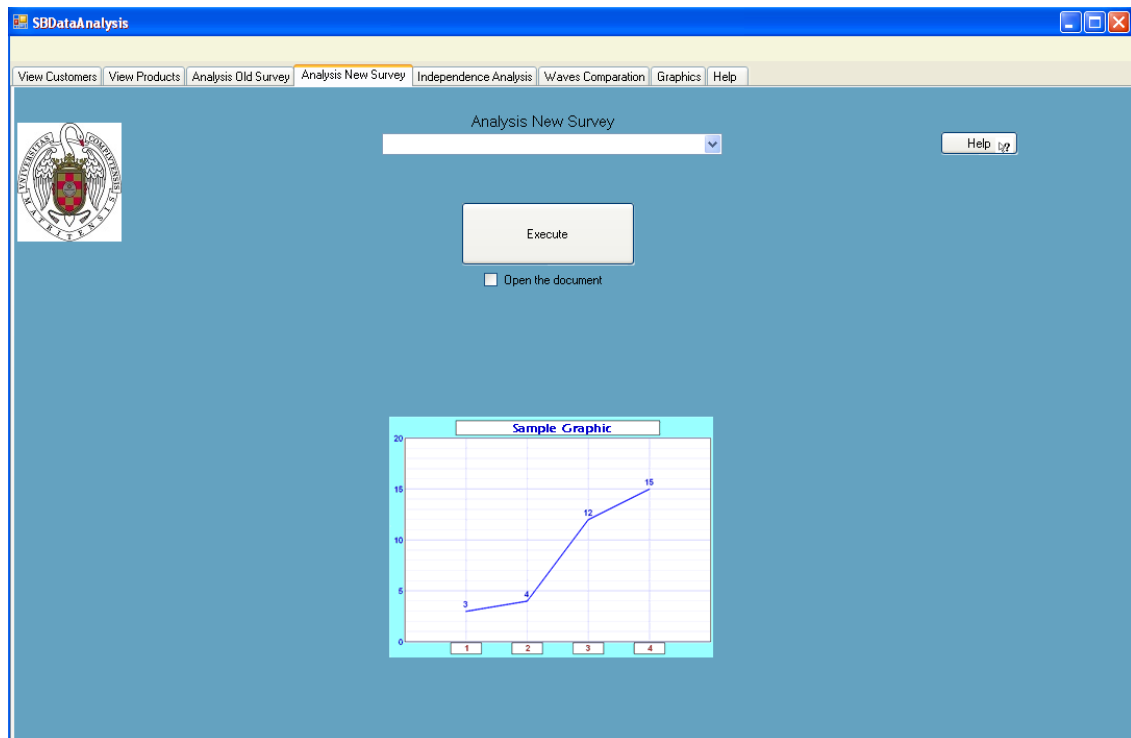


Figura 35 Al pulsar sobre Help se genera la ayuda

Una vez pulsado el botón, la aplicación abre la pestaña Help. Se observa en la figura 36 el texto explicativo de ayuda para manejar la pestaña así como un botón back que devuelve a la pestaña desde la cual hemos solicitado la ayuda.

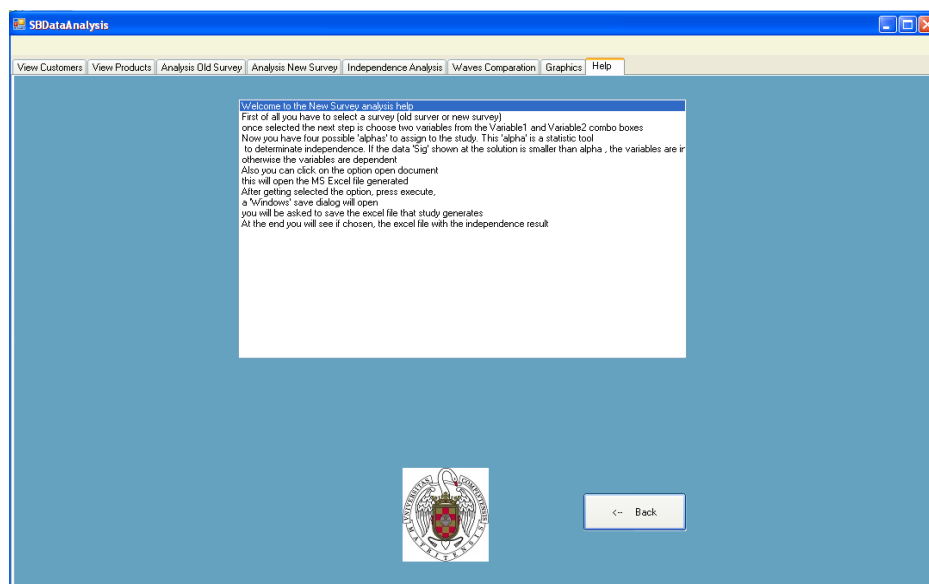
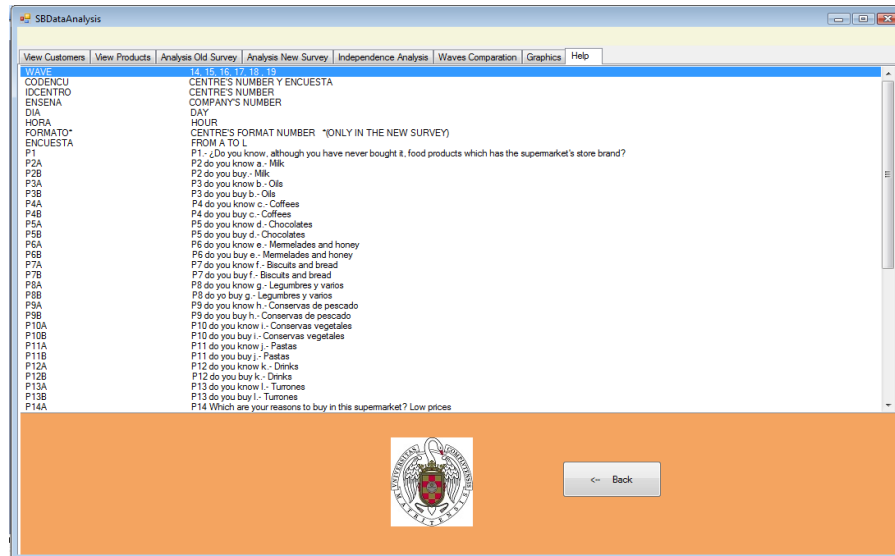


Figura 35. Pestaña Help con el texto explicativo de la pestaña New Survey Analysis

En el caso de las pestañas de visualización al apretar sobre el botón *help* se mostrará en la pestaña Help el significado de las diferentes variables que forman la B.D. En la figura 37 se puede observar la pestaña help con el significado de las variables de las encuestas a los consumidores.



**Figura 37 .** Pestaña Help con el significado de las diferentes variables de la encuesta de consumidores



## 4. Manual de uso

Para la creación de la manual de uso de SBDataAnalysis se ha utilizado una herramienta de autor, TT Knowledge Force, que ha permitido la creación tanto de la misma manual de uso, como de un documento e-Learning. El documento e-Learning permite al usuario de una forma dinámica y amena aprender a usar el programa. Muestra cómo se hace cada acción y hace participe de la realización de estas acciones al usuario, que aprende de una forma más rápida a usar la aplicación.

Para aprender a usar el software TTKnowledge Force se ha asistido a un curso impartido en la Facultad de Informática de la Universidad Complutense de Madrid, en la que se explicaba el uso de la herramienta de autor para la creación tanto de la manual de uso, como del documento e-Learning.

### 4.1. Herramientas de autor

Las herramientas de autor o metamedios son aplicaciones que permiten un trabajo multimedia y constructivo para generar un entorno de aprendizaje dinámico. Dentro de las funcionalidades que este tipo de herramientas presentan se puede destacar la posibilidad de crear actividades o pequeñas aplicaciones desde la misma herramienta.

Las herramientas de autor tienen distintos módulos desde los cuáles se pueden organizar actividades o se pueden interconectar pequeños componentes y se pueden adecuar a los objetivos, los conocimientos y habilidades que se busque desarrollar por parte del autor.

Las herramientas de autor más básicas son aquellas que solamente permiten un conjunto limitado de acciones para que el usuario interactúe con el sistema, como por ejemplo, crear mapas sensibles o ir de una diapositiva a la siguiente. Ejemplos de estas herramientas pueden ser kiss u OpenOffice.org Impress o TT Knowledge Force. Las herramientas más avanzadas incluyen lenguajes de programación o la creación de ambientes con flash.

En general todas estas herramientas permiten la creación de documentos e-learning, o aprendizaje electrónico; de documento en formato HTML, de presentaciones y de manuales de uso en formato Word o PDF.

El **e-learning** es una forma de enseñanza a través de correo electrónico o a distancia en el que se integra el uso de las tecnologías de la información y otros elementos pedagógicos (didácticos) para la formación, capacitación y enseñanza de los usuarios o estudiantes en línea, es decir, se puede entender como una modalidad de aprendizaje dentro de la educación a distancia. El e-learning utiliza herramientas y medios diversos como Internet, intranets, CD-ROM, producciones multimedia (Textos, imágenes, audio, video, etc.), entre otros.

E-learning es principalmente un medio electrónico para el aprendizaje a distancia o virtual, donde se puede interactuar con los profesores por medio de Internet. El usuario puede manejar los horarios, es un medio completamente autónomo. Constituye una propuesta de formación que contempla su implementación predominantemente mediante Internet, haciendo uso de los servicios y herramientas que esta tecnología provee.

Dentro de la modalidad de educación a distancia, el e-learning es una de las opciones que actualmente se utiliza con mayor frecuencia para atender la necesidad de educación continua o permanente. La generación de programas de perfeccionamiento profesional no reglados está en crecimiento debido a que existe un reconocimiento de que los trabajadores se capaciten y se adapten a los nuevos requerimientos productivos. El e-learning, dadas sus características y el soporte tecnológico que lo respalda, se constituye en una alternativa para aquellos que combinan trabajo y estudio, ya que no es necesario acudir a una aula permanentemente.

Si la educación a distancia es, desde sus orígenes, una opción para atender la formación de personas, el e-learning tiene la ventaja de que los usuarios eligen sus propios horarios, y puede entrar a la plataforma desde cualquier lugar donde puedan acceder a un ordenador que tenga conexión a Internet. La educación virtual da la oportunidad de que el estudiante elija sus horarios de estudio convirtiéndose así en una muy buena opción para aquellas personas autónomas que trabajen y quieran estudiar en sus momentos libres; por otra parte es importante mencionar que el e-learning es una excelente herramienta que puede ayudar a los usuarios no solo a aprender conceptos nuevos sino también a afianzar conocimientos y habilidades, aumentando así la autonomía y la motivación de los estudiantes por diferentes temas.

La herramienta **TT Knowledge Force** es una aplicación de software que sirve de soporte a los procesos de conocimiento y a la transferencia de conocimiento en el entorno empresarial. Permite estructurar modelos de proceso y de conocimiento integrados. En lo que respecta a la elaboración y actualización de modelos, TT Knowledge Force facilita la colaboración de todos los miembros de un equipo de proyecto.

Las funciones principales de TT Knowledge Force son la configuración y actualización de los medios Documentación y e-Learning interactivo. TT Knowledge Force permite estructurar los contenidos en cursos y temas y la traducción cuasi-automática de sus contenidos a otros idiomas.

En TT Knowledge Force los elementos de conocimiento se presentan en forma de documentos. Técnicamente, éstos constituyen una unidad cerrada. Los documentos pueden tener formatos de archivo existentes, tales como \*.doc, \*.ppt, \*.xls o \*.html. Estos archivos se pueden importar y administrar en TT Knowledge Force. No obstante, TT Knowledge Force utiliza el documento TT para elaborar documentación y e-Learning de forma eficiente. Se trata de un formato de origen exclusivo en el que se pueden almacenar y actualizar sin redundancias los medios Documentación y e-Learning.

Todos los documentos de TT Knowledge Force poseen metadatos. Éstos sirven para identificar de forma única y describir con mayor detalle (p. ej. el título de un documento). Los dos medios que constituyen un documento TT son: La documentación y e-Learning. Ambos se pueden actualizar y exportar o publicar en un documento TT. La aplicación graba el uso de la aplicación de la que se quiere hacer el documento TT y, a partir de ahí, crea automáticamente los dos medios.

E-Learning es un medio de aprendizaje digital basado en HTML y orientado a páginas. Se compone de páginas consecutivas y siempre incluye un elemento de navegación (la ventana de comentario) mediante el cual se puede alternar entre páginas. El medio e-learning se presenta en dos modos: modo de aprendizaje y modo examen.

En el modo Aprendizaje el usuario elabora el contenido de un objeto de conocimiento. Dentro de este modo existen tres modos de reproducción: Modo interactivo, modo presentación y modo Película. El modo Examen le presenta al usuario una situación de examen en la que se trabaja con preguntas y secciones de simulaciones. El resultado se registra y puede verse en la página de evaluación.

El medio de documentación describe el contenido a trasladar en forma de texto de flujo. Este texto puede estar disponible en formato digital o impreso. TT Knowledge Force permite, por ejemplo, generar características de la documentación para un documento TT.

De todas las posibles funcionalidades proporcionadas por el programa TT Knowledge Force, para la creación del manual de uso y del documento de e-Learning se ha usado la documentación y el modo de aprendizaje respectivamente. Con el uso del software TT Knowledge Force se ha podido crear a la vez los dos tipos de medios, tanto el documento con la manual de uso, como el documento e-Learning.

El medio de documentación permite la creación de la manual de uso una forma automática y fácil a partir de la grabación que realiza de la aplicación. La documentación generada es de fácil comprensión ya que genera comentarios y señala en las imágenes los pasos a seguir. Con esto, el usuario no necesita una larga explicación, porque puede ver todos los pasos en las imágenes creadas.

En la figura 38 se puede observar la documentación que la herramienta genera automáticamente. La herramienta señala los pasos a realizar con un número y un rectángulo, que señala el objeto sobre el que se realiza la acción.

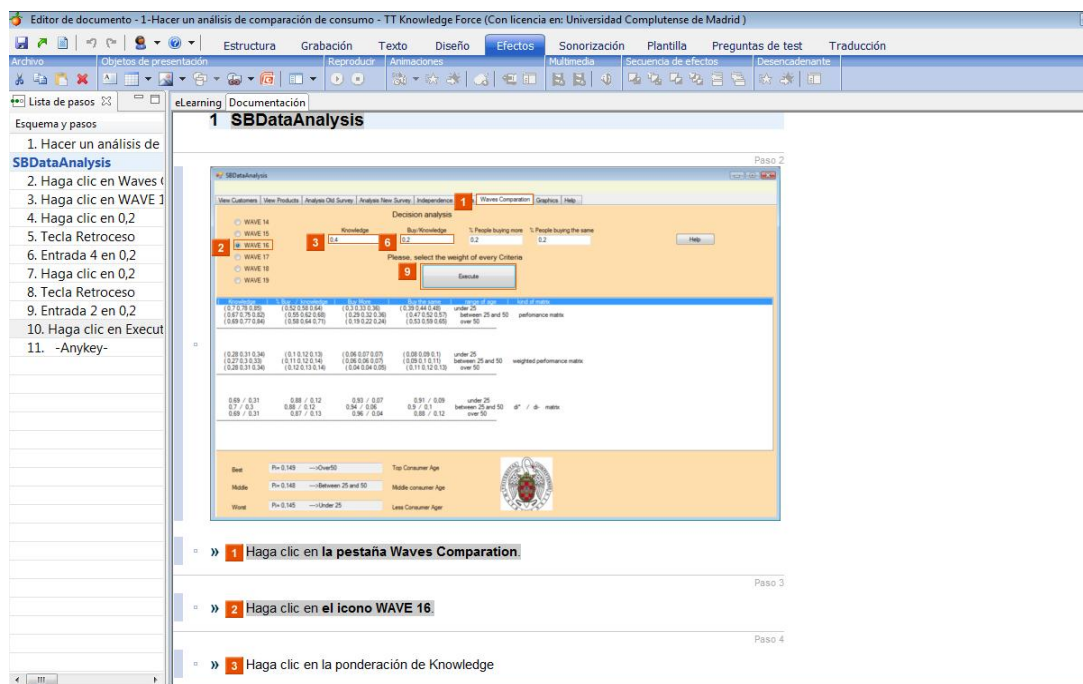


Figura 38. Documentación generada automáticamente por la herramienta

El medio e-Learning, y en este caso el modo de aprendizaje, permite la creación del documento interactivo con el que el usuario puede practicar y simular el uso de la aplicación. Esta simulación es creada automáticamente a partir de la grabación realizada, por lo que el usuario tiene la sensación de usar realmente la aplicación.

Además de hacer una simulación de la aplicación, se puede añadir elementos de presentación para hacer mucho más entendible y lustroso las acciones que se quieren enseñar al usuario. En la figura 39 se puede observar el documento e-Learning que la herramienta genera automáticamente. La herramienta genera la simulación e incluye una tabla de navegación en dónde se indica los pasos a realizar y permite al usuario manejar la simulación.



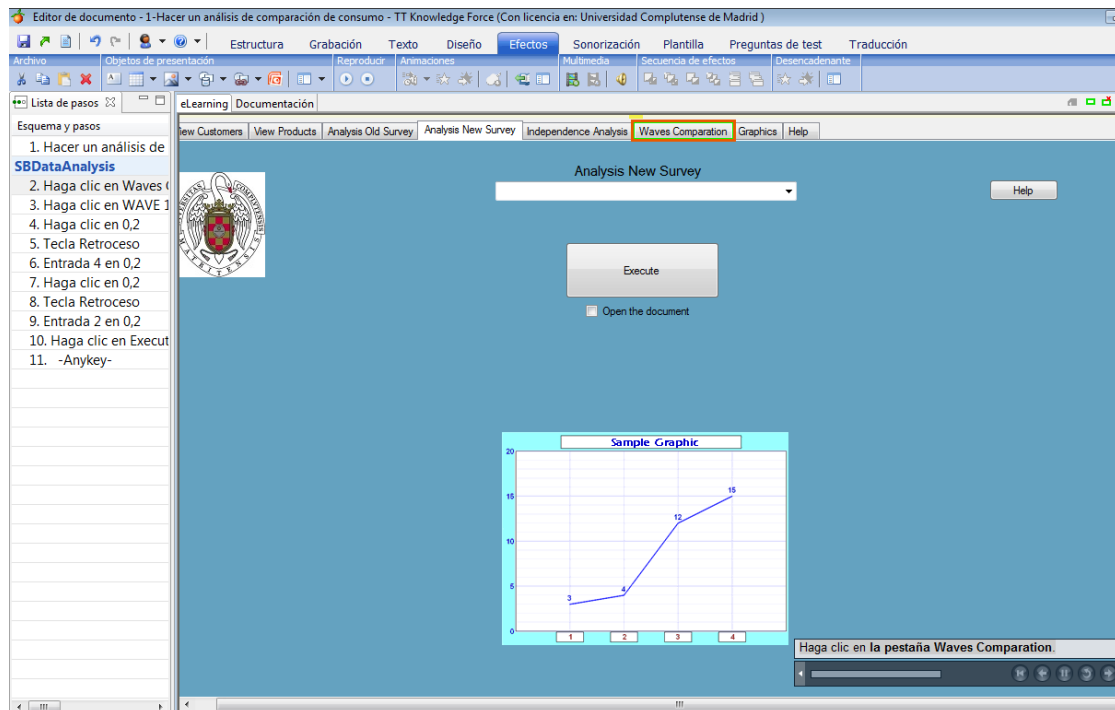


Figura 39. simulación de e-Learning generada automáticamente por la herramienta

En la figura 40 se puede observar cómo se le puede añadir a una simulación elementos de presentación, en este caso para explicar los diferentes componentes de la pestaña. A estos elementos de presentación se le puede añadir efectos, con esto la simulación es mucho más visual y atractiva para el usuario.

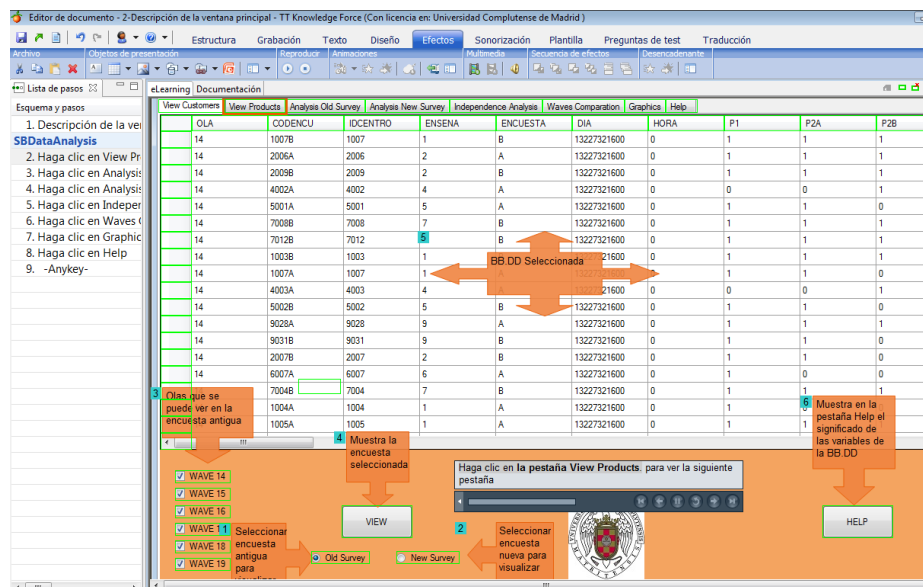


Figura 40. Simulación generada con elementos de presentación que explican las partes de la pestaña

## **4.2. Manual de uso y documento e-Learning**

El manual de uso se encuentra en el Anexo 1 del documento.

El documento e-Learning interactivo se encuentra en el CD adjunto al documento.

## 5. Conclusiones y trabajo futuro

Con la aplicación desarrollada se ha conseguido un software *open source* para realizar numerosos estudios de encuestas a consumidores sobre marcas de distribuidor. Estos estudios ayudarán al Departamento de Investigación y Comercialización de Mercados de la Universidad Complutense de Madrid a sacar conclusiones con respecto a los gustos y tendencias de los consumidores referido a las marcas de distribuidor.

En los estudios de comparación de consumo por rangos de edad y olas se ha aplicado técnicas de Soft-Computing, metodología de investigación actualmente en auge (en particular técnicas de decisión multicriterio), que ha permitido realizar el estudio. Sin esta técnica, la realización del estudio hubiera sido más difícil y con un resultado no tan exacto y fiable.

Se ha realizado una aplicación en la que los resultados de los estudios se expresan en documentos claros e intuitivos, fácilmente interpretables y que permiten el transporte de los datos a otros medios.

Actualmente, el mercado del software dedicado a hacer estudios estadísticos está formado por programas con un interfaz que puede llegar a ser complejo, en la que se necesita mucho tiempo para aprender a usar todas sus opciones. Esta aplicación tiene un interfaz amigable, sencillo de usar en la que se recoge todas las necesidades que el Departamento de Investigación y Comercialización de Mercados expreso en las diferentes reuniones realizadas. La aplicación está pensada para que el usuario sepa, en todo momento, cómo se realizan todas las acciones disponibles; además, todas las acciones están pensadas para hacerse de una manera sencilla, intuitiva y rápida.

La realización de este proyecto, en especial su apartado relacionado con el Soft-Computing ha dado lugar a la elaboración de varios artículos que han sido aceptados en diversos congresos, algunos tan importantes como el Congreso Español de Informática (CEDI), que se celebrará en Valencia entre los días 7 y 10 de septiembre. Así como en el XXII Congreso Nacional de Estadística e Investigación Operativa, que se celebra en La Coruña entre el 14 y 17 de Septiembre.

El sistema desarrollado puede ser ampliado en diversos puntos, mejorando sus funciones y eficacia. El número de análisis que se realizan son limitados, se podría aumentar el número de análisis realizados. Una manera sería hablar con el Departamento de Investigación y Comercialización de Mercados, que podría sugerir nuevos estudios, que serían útiles para ayudarles a realizar nuevas investigaciones en el mundo de las marcas de distribuidor. Otra

manera, más completa, sería que fuera el usuario el que eligiera el análisis que quiere realizar, que variables intervienen y qué estadísticos quiere obtener como resultado.

El análisis de independencia se realiza sobre dos variables, una posible ampliación sería aumentar el número de variables que interviene en el estudio, hacer un estudio de independencia entre tres o cuatro variables, según quiera el usuario.

Una ampliación muy interesante sería aplicar las técnicas de decisión multicriterio, en otros estudios que el usuario pudiera elegir tanto las variables a tener en cuenta como los criterios y sus ponderaciones o pesos.

Los gráficos que se muestran con los resultados, muestran el resultado para la media aritmética. Otra ampliación sería la de mostrar gráficos con los resultados de otros estadísticos que el usuario eligiera, desviación típica, varianza, media geométrica etc.

Por último, una ampliación que ayudaría al usuario a comprender y usar mejor la herramienta sería hacer una guía de uso, usando una herramienta de autor, en formato e-Learning en el que, además de mostrarle el uso de la aplicación, se le permita hacer ejercicios y exámenes. Con estos ejercicios y exámenes el usuario aprendería más rápido y mejor el uso de la herramienta y a interpretar los resultados de los análisis, además de responder de forma autodidacta dudas que tenga sobre la herramienta

## 6. Bibliografía

- Gómez Villegas, M. A. Inferencia Estadística, Ed. Diaz de Santos, I.S.B.N : 84-7978-687-6, 2005.
- Huifang Mao y H. Sanker Krishnan. Rrects of Prototype and Exemplar Fit on Brand Extension Evaluations: A two-Process Contingency Model. Journal of Consumer Research, Vol. 33, pp: 41-49, 2006.
- <http://msdn.microsoft.com/es-es/vstudio/default.aspx>.
- [Puelles, J. A. Análisis del fenómeno de las m arcas de distribuidor en España y de su tratamiento estratégico. Boletín ICE. Comercio Interior: Regulación y Estrategias Empresariales, nº 739, pp: 117-129, 1995.
- Puelles, J. A. El canal de distribución de la industria alimentaria española: reflexiones sobre la evolución del entorno que determina importantes cambios en el mismo. Alta Dirección, nº 139, pp:169-176, 1988.]
- Puelles, J. A. y Puelles, M. Marcas de Distribuidor. Distribución y Consumo. Pp: 55-71 , 2003.
- PuellesJ.A y Puelles, M. Marcas de Distribuidor (MDD). 100 ideas clave. Distribución y Consumo Pp 241-256, 2008
- Wei Wang y Norman Fenton, %Risk and Confidence Analysis for Fuzzy Multicriteria Decision Making+, Risk Analysis and Decision Research Group, pp. 1-7.
- Manual de uso de TT Knowledge Force Pp:17-19, 2009
- *Curso de SPSS*  
<<http://www.spssfree.com/>> [
- SPSS  
<[www.spss.es](http://www.spss.es)>
- Support SPSS  
<<http://support.spss.com/>>
- *SPSS devcentral*  
<<http://www.spss.com/devcentral>>
- *Foro de SPSS*  
<<http://www.spss.com/fusetalk/forum/>>
- *Wikipedia*  
<<http://es.wikipedia.org/>>

- *Licencia pública general GNU (GNU GPL).*  
<<http://www.viti.es/gnu/licenses/gpl.html/>> [consulta Mayo 2010]
- Nielsen  
<[www.nielsen.com](http://www.nielsen.com)>
- Marketing UCM  
<<http://www.ucm.es/info/marketing/>>
- Pruebas chi-cuadrado ajuste de independencia  
[http://www.ucm.es/info/genetica/Estadistica/estadistica\\_basica%202.htm#Pruebas%20chi-cuadrado%20de%20ajuste%20e%20independencia](http://www.ucm.es/info/genetica/Estadistica/estadistica_basica%202.htm#Pruebas%20chi-cuadrado%20de%20ajuste%20e%20independencia)
- Marketing comunidad  
<[www.marketingcomunidad.com](http://www.marketingcomunidad.com)>
- Lotto Wizard  
<<http://www.lotto-wizard.net/>>
- EPIDAT  
<<http://www.paho.org/Spanish/SHA/epidat.htm>>

## 7. Anexo

Se adjuntan, a modo de anexo, los siguientes documentos:

- **Anexo 1:** Manual de uso SBDataAnalysis.
- **Anexo 2a:** Guía de investigación sobre Marcas de distribuidor.
- **Anexo 2b:** Encuesta a consumidores.







**UNIVERSIDAD COMPLUTENSE DE MADRID**



**FACULTAD DE INFORMÁTICA**

**SISTEMAS INFORMÁTICOS**

**CURSO 2009/2010**

**ANEXO 1**

**MANUAL DE USO**

**BSDataAnalysis**

**Jorge Caballero Prados**

**Pablo Rodrigo Jimeno Hidalgo**

**Raúl Ortega Maestro**

*Dirigido por:*

***Dra. María Victoria López López***

*Departamento de Arquitectura de Computadores y Automática*

*Facultad de Informática - Universidad Complutense de Madrid*



# Índice

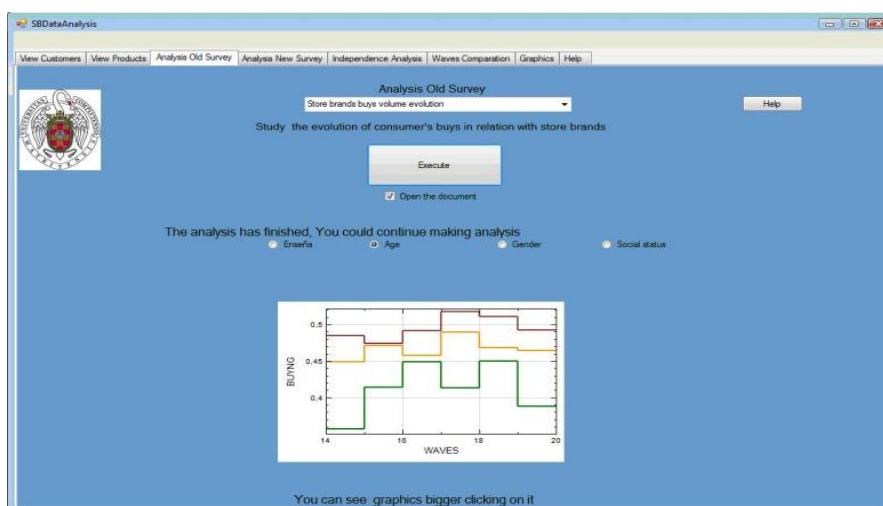
|         |   |     |
|---------|---|-----|
| 1       | Introducción a SBDataAnalysis .....                               | 105 |
| 1.1     | ¿Qué es SBDataAnalysis? .....                                     | 105 |
| 1.2     | Instalación, requisitos y versiones de SBDataAnalysis.....        | 106 |
| 1.2.1   | Requisitos para la aplicación.....                                | 106 |
| 1.2.2   | Versiones de SBDataAnalysis.....                                  | 106 |
| 1.2.3   | Instalación de la versión del usuario .....                       | 109 |
| 1.2.4   | Cómo se ejecuta SBDataAnalysis .....                              | 110 |
| 1.3     | Descripción de ventanas .....                                     | 113 |
| 1.3.1   | Descripción de la portada .....                                   | 113 |
| 1.3.2   | Descripción de la ventana principal .....                         | 114 |
| 1.3.2.1 | View Costumers .....  | 115 |
| 1.3.2.2 | View Products .....   | 115 |
| 1.3.2.3 | Analysis Old Survey .....   | 116 |
| 1.3.2.4 | Analysis New Survey.....  | 116 |
| 1.3.2.5 | Independence Analysis .....                                       | 117 |
| 1.3.2.6 | Waves Comparation.....  | 117 |
| 1.3.2.7 | Graphics .....  | 118 |
| 1.3.2.8 | Help.....   | 118 |
| 2       | Visualización de las Bases de Datos.....                          | 119 |
| 2.1     | Visualización de BB.DD encuesta consumidores.....                 | 119 |
| 2.1.1   | Visualización encuesta antigua .....                              | 119 |
| 2.1.2   | Visualización encuesta nueva .....                                | 121 |
| 2.1.3   | Ver significado de las variables de consumidores.....             | 123 |
| 2.2     | Visualización de BB.DD de productos de MDD .....                  | 124 |
| 2.2.1   | Visualización de una Ola de productos .....                       | 124 |
| 2.2.2   | Ver significado de las variables de los productos .....           | 126 |
| 3       | Análisis de encuestas de consumidores .....                       | 127 |
| 3.1     | Análisis de encuesta de consumidores antigua .....                | 127 |
| 3.1.1   | Hacer un análisis de las encuestas antiguas a consumidores .....  | 127 |
| 3.1.2   | Ampliar gráfico de resultado.....                                 | 131 |
| 3.1.3   | Ver ayuda de la pestaña .....                                     | 132 |
| 3.2     | Análisis de encuesta de consumidores nueva .....                  | 134 |
| 3.2.1   | Hacer un análisis de las encuestas nuevas a los consumidores..... | 134 |
| 3.2.2   | Ampliar gráfico de resultado.....                                 | 137 |
| 3.2.3   | Ver ayuda de la pestaña .....                                     | 138 |
| 4       | Análisis de independencia.....                                    | 141 |
| 4.1     | Hacer un análisis de independencia .....                          | 141 |
| 4.2     | Ver ayuda de la pestaña .....                                     | 145 |
| 5       | Comparación de consumo .....                                      | 147 |
| 5.1     | Hacer un análisis de comparación de consumo .....                 | 147 |
| 5.2     | Ver ayuda de la pestaña .....                                     | 148 |



# 1 Introducción a SBDataAnalysis

## 1.1 ¿Qué es SBDataAnalysis?

- SBDataAnalysis es una aplicación open-source que facilita realización de estudios sobre unas bases de datos de encuestas de Marcas de Distribuidor.
- Permite la ejecución de diferentes estudios sobre la evolución de las respuestas, la comparación de variables por enseñas o formatos de supermercados.
- Permite el estudio de independencia entre las respuestas de las encuestas.
- Permite la comparación de consumo por rango de edad y por olas.
- Permite ver las bases de datos y ver gráficos de los estudios que se realizan de evolución y comparación de variables por enseñas o formatos de supermercados.
- Todos los estudios de evolución, comparación de variables por enseñas o formatos de supermercados y de estudio de independencia se crean y guardan en un archivo Excel bien ordenado y con todos los resultados de los estudios.



Microsoft Excel - prueba compra

Archivo Edición Ver Insertar Formato Herramientas Datos Ventana ?

Escribe una pregunta

| A1 | A                   | B                       | C           | D        | E             | F              | G        | H              | I       |
|----|---------------------|-------------------------|-------------|----------|---------------|----------------|----------|----------------|---------|
| 16 |                     | Elapsed Time            | 0:00:00.125 |          |               |                |          |                |         |
| 17 |                     |                         |             |          |               |                |          |                |         |
| 18 |                     | Case Processing Summary |             |          |               |                |          |                |         |
| 19 |                     |                         | Cases       |          |               |                |          |                |         |
| 20 |                     |                         | Included    | Excluded |               | Total          |          |                |         |
| 21 |                     |                         | Percent     | N        | Percent       | N              | Percent  |                |         |
| 22 | COMPRA * OLA * EDAD | 3898                    | 100,0%      | 0        | 0%            | 3898           | 100,0%   |                |         |
| 23 |                     |                         |             |          |               |                |          |                |         |
| 24 |                     | Report                  |             |          |               |                |          |                |         |
| 25 | COMPRA              |                         |             |          |               |                |          |                |         |
| 26 | OLA                 | EDAD                    | N           | Mean     | Harmonic Mean | Geometric Mean | Variance | Std. Deviation | Minimum |
| 27 | 14                  | menor de 25 años        | 10          | ,3583    |               | ,0000          | ,059     | ,24232         | ,00     |
| 28 |                     | entre 25 y 50 años      | 61          | ,4495    |               | ,0000          | ,059     | ,24356         | ,00     |
| 29 |                     | mayor de 50 años        | 35          | ,4857    |               | ,0000          | ,058     | ,24126         | ,00     |
| 30 |                     | Total                   | 106         | ,4529    |               | ,0000          | ,059     | ,24289         | ,00     |
| 31 | 15                  | menor de 25 años        | 57          | ,4152    |               | ,0000          | ,040     | ,20013         | ,00     |
| 32 |                     | entre 25 y 50 años      | 230         | ,4725    |               | ,0000          | ,076     | ,27517         | ,00     |
| 33 |                     | mayor de 50 años        | 134         | ,4751    |               | ,0000          | ,079     | ,28135         | ,00     |
| 34 |                     | Total                   | 424         | ,4652    | ,3529         | ,3816          | ,049     | ,22048         | ,25     |
| 35 |                     | Total                   | 424         | ,4652    |               | ,0000          | ,072     | ,26801         | ,00     |
| 36 | 16                  | menor de 25 años        | 94          | ,4495    |               | ,0000          | ,056     | ,23640         | ,00     |
| 37 |                     | entre 25 y 50 años      | 337         | ,4585    |               | ,0000          | ,065     | ,25476         | ,00     |
| 38 |                     | mayor de 50 años        | 194         | ,4923    |               | ,0000          | ,062     | ,24880         | ,00     |
| 39 |                     | Total                   | 625         | ,4676    |               | ,0000          | ,063     | ,25041         | ,00     |
| 40 | 17                  | menor de 25 años        | 102         | ,4142    |               | ,0000          | ,046     | ,21478         | ,00     |
| 41 |                     | entre 25 y 50 años      | 359         | ,4905    |               | ,0000          | ,071     | ,26675         | ,00     |
| 42 |                     | mayor de 50 años        | 177         | ,5184    |               | ,0000          | ,065     | ,25550         | ,00     |
| 43 |                     | Total                   | 638         | ,4742    |               | ,0000          | ,078     | ,24150         | ,00     |
| 44 |                     |                         |             |          |               |                |          |                |         |

Sheet1

Listo

## 1.2 Instalación, requisitos y versiones de SBDataAnalysis

### 1.2.1 Requisitos para la aplicación

La aplicación SBDataAnalysis tiene como requisitos básicos:

- **SPSS Versión 17.**
- **SPSSStatisticsNetPlugInWin32\_1700.** PlugIn que permite la conexión entre la aplicación SBDataAnalysis y SPSS.
  - El plugIn se adjunta en las carpetas de la aplicación pero se puede descargar de: **¡Error! Referencia de hipervínculo no válida.**

En el caso de los desarrolladores. La aplicación ha sido desarrollada sobre **C#** en *Visual Studio 2008*.

### 1.2.2 Versiones de SBDataAnalysis

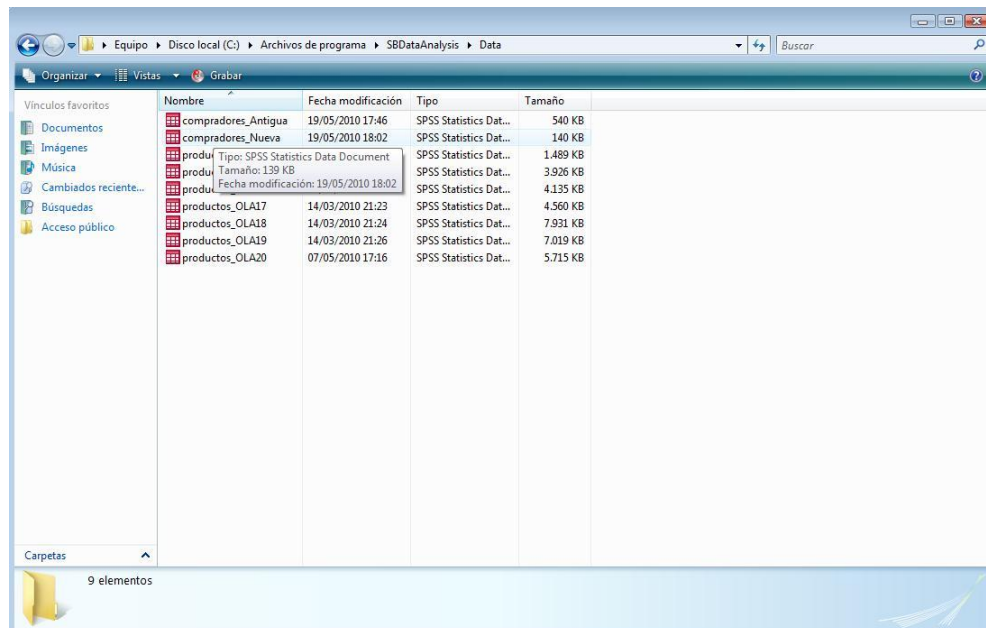
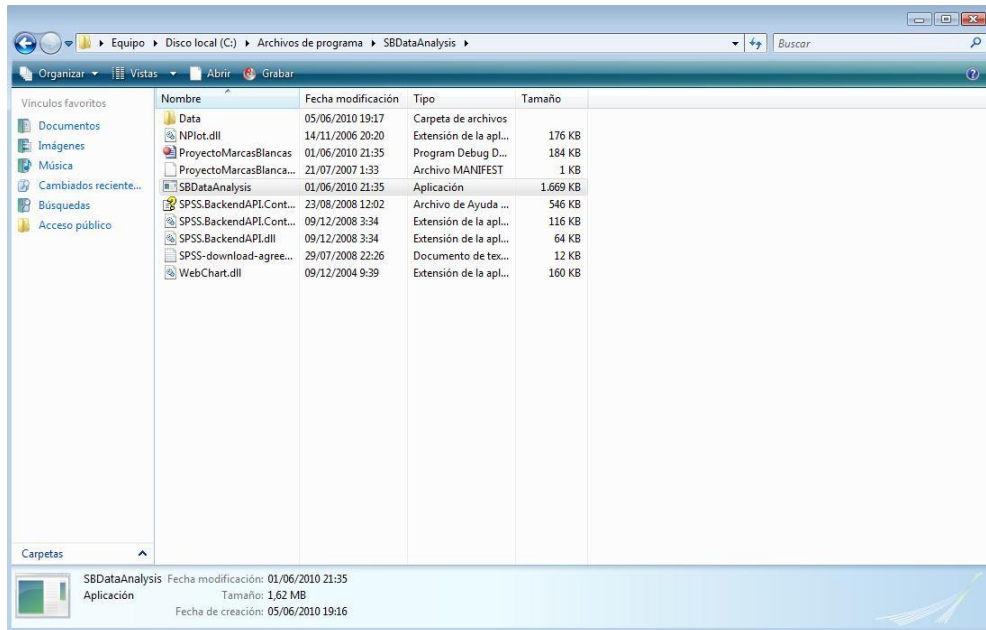
Existen dos versiones de SBDataAnalysis:

- Versión para usuario, en la que se encuentra la aplicación.
- Versión para desarrolladores, en la que se encuentra el programa en C# .

La versión de usuario se encuentra dentro del archivo comprimido *SBDataANalysis\_User.rar*. En el archivo comprimido se encuentra la carpeta *SBDataAnalysis*. Los principales componentes de esta carpeta son:

- El ejecutable de la aplicación compilado y listo para usar, cuyo nombre es *SBDataAnalysis*.
- Carpeta Data. Dentro de esta carpeta se encuentran las diferentes BB.DD que se usan en la aplicación.

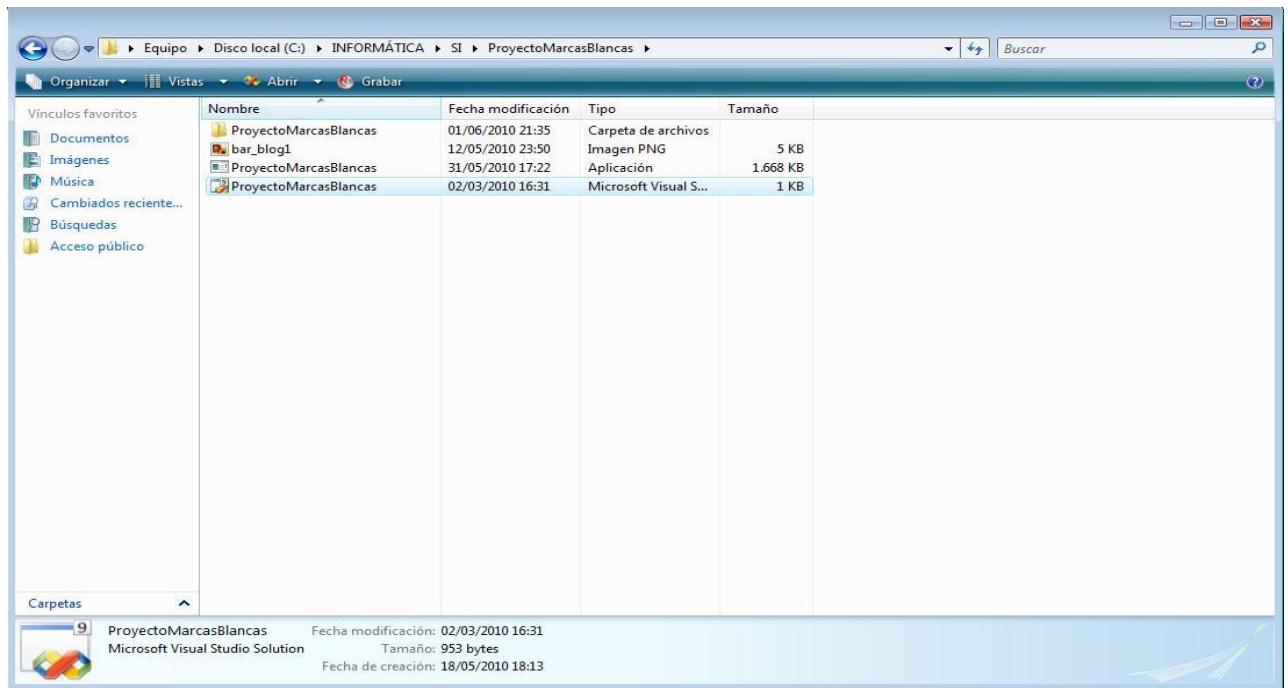
Además de la carpeta *SBDataAnalysis*, en el archivo *SBDaAnalysis\_User.rar* se encuentra el plugIn *SPSSStatisticsNetPlugInWin32\_1700*



La carpeta SBDDataAnalysis debe instalarse en **C:/Archivos de programa** para que funcione la aplicación.

La versión para desarrolladores se encuentra dentro del archivo comprimido *SBDDataAnalysis\_Program*. En el archivo comprimido se encuentra la carpeta *ProyectoMarcasBlancas*. Los principales componentes de esta carpeta son:

- El archivo de Microsoft Visual Studio Solution *ProyectoMarcasBlancas* que es la cabecera del proyecto de la aplicación.
- La Carpeta *ProyectoMarcasBlancas*. Dentro se encuentran las diferentes clases y recursos de la aplicación.
- Esta versión se puede instalar dónde se desee.

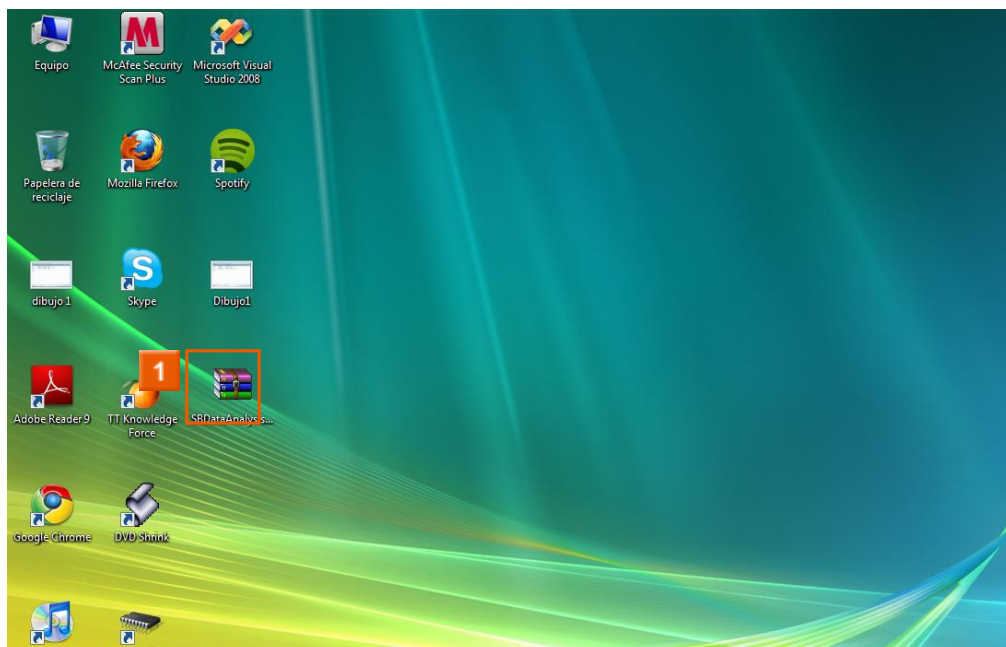


⚠ Además de esta carpeta se debe guardar la carpeta de la versión del usuario. Esa carpeta, *SBDDataAnalysis*, debe instalarse en *C:/Archivos de programa* para que funcione la aplicación y el programa.

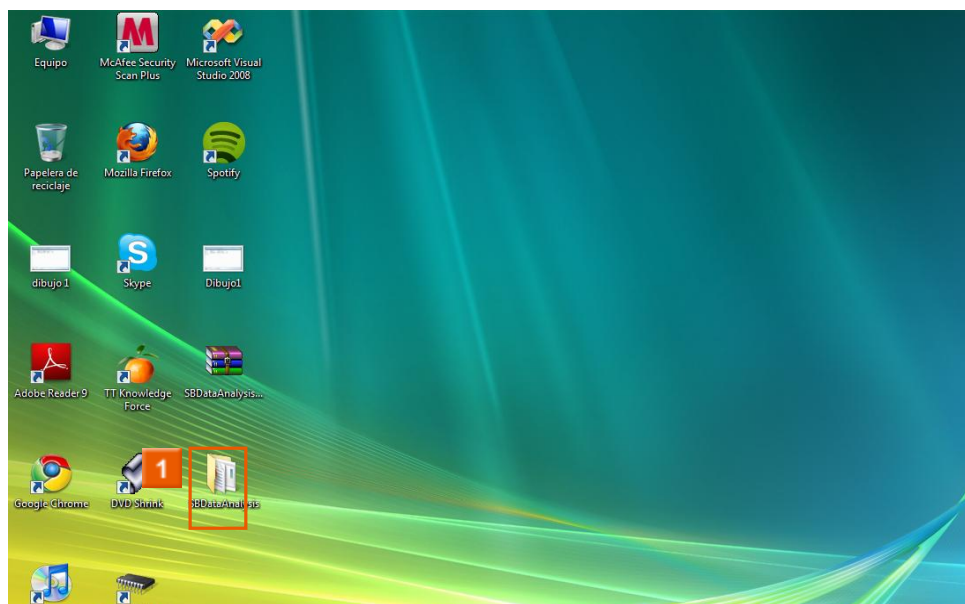


### 1.2.3 Instalación de la versión del usuario

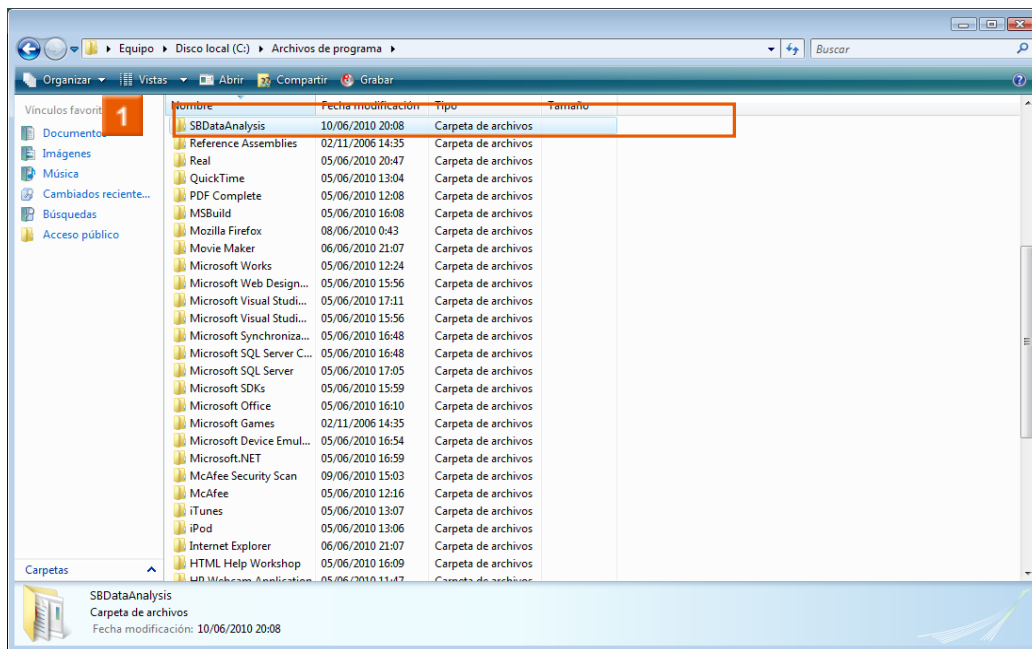
A continuación se describe cómo instalar la versión del usuario.



**1** Haga clic en el archivo comprimido **SBDDataAnalysis\_User**. A continuación extraiga el contenido del archivo.

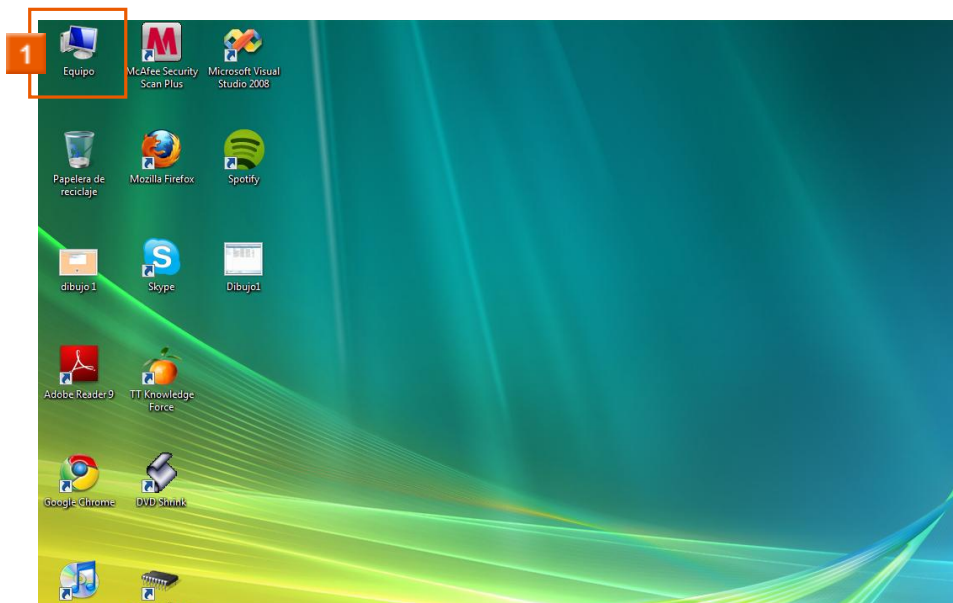


**1** Seleccione el documento SBDDataAnalysis, y cópielo en C:/Archivos de programa.

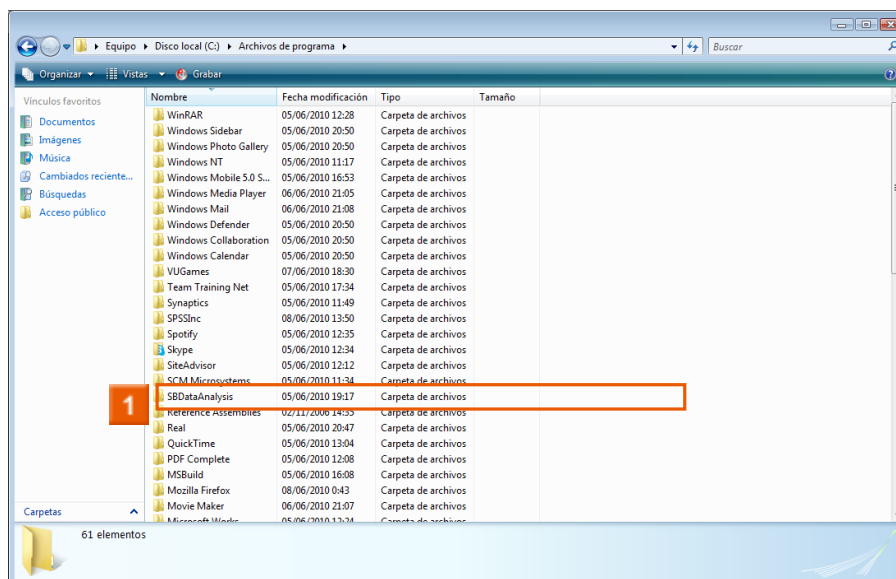


#### 1.2.4 Cómo se ejecuta SBDataAnalysis

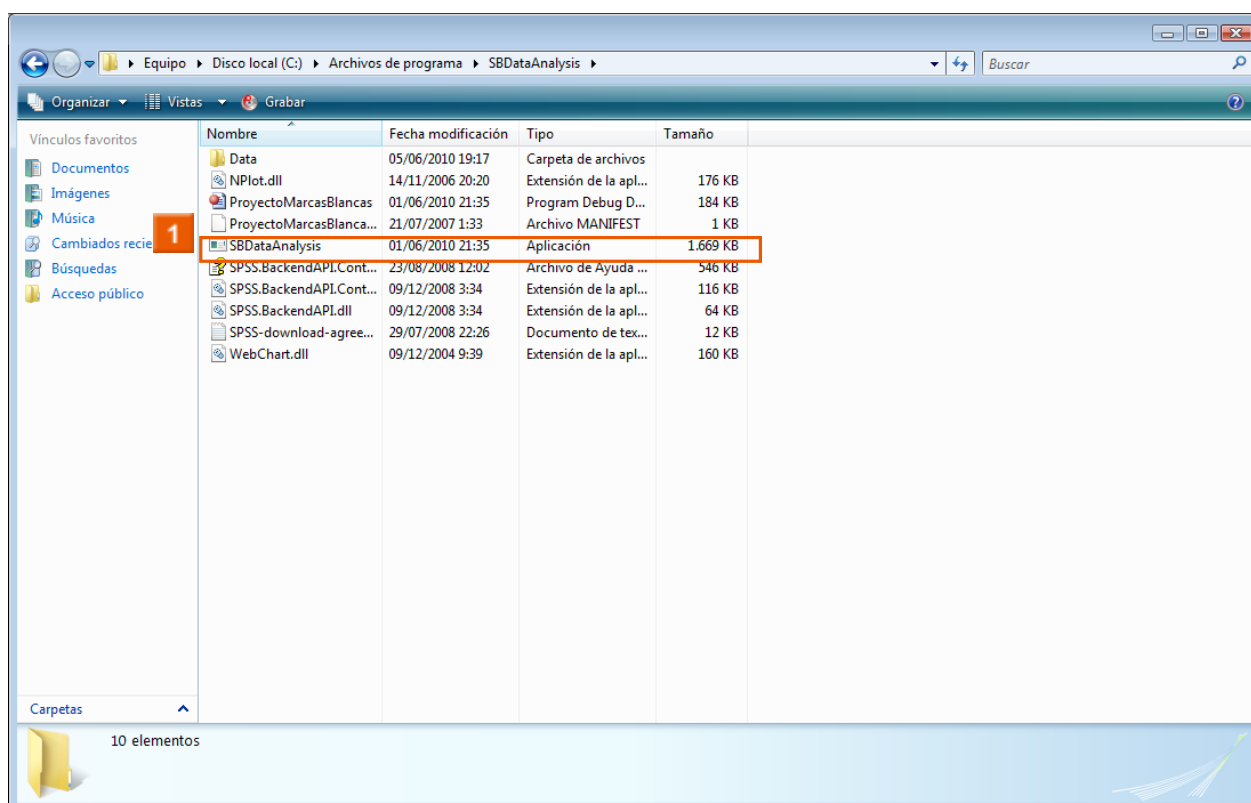
A continuación se describe cómo ejecutar el programa SBDataAnalysis. El programa debe encontrarse en el sitio dónde se ha especificado en el apartado **Instalación de SBDataAnalysis**.



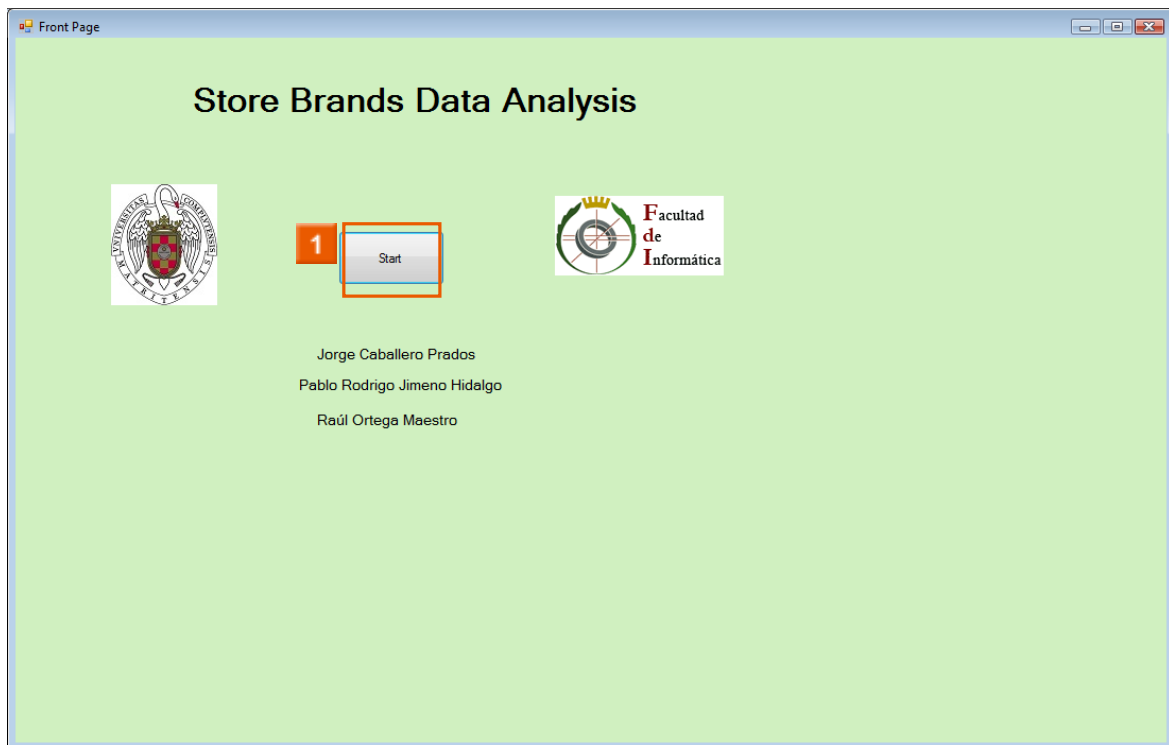
**1** Haga doble clic en **el elemento de lista Equipo** y diríjase a C:/Archivos de programa.



1 Haga doble clic en el elemento de lista **SBDataAnalysis**.



1 Haga doble clic en el ejecutable **SBDataAnalysis** para abrir la aplicación.



**1** Haga clic en **Start** para ir a la ventana principal de la aplicación.

## 1.3 Descripción de ventanas

### 1.3.1 Descripción de la portada

A continuación se describe la portada de la aplicación.



### 1.3.2 Descripción de la ventana principal

A continuación se muestra una descripción de la ventana principal y sus pestañas. Se mostrará para qué sirve cada elemento de las pestañas.

Existen 8 pestañas cuyos nombres y funciones son:

**View Costumers:** Pestaña en la que muestra las distintas BB.DD de encuestas a consumidores.

**View Products:** Pestaña en la que muestra las distintas BB.DD de los productos de MDD.

**Analysis Old Survey:** Pestaña en la que realiza análisis sobre las encuestas antiguas a consumidores.

**Analysis New Survey:** Pestaña en la que realiza análisis sobre las encuestas nuevas a consumidores.

**Independece Analysis;** Pestaña en la que realiza un análisis de independencia sobre dos variables con respecto a un tipo de encuesta de consumidores, antigua o nueva.

**Waves Comparison:** Pestaña en la que realiza un análisis de consumo por rango de edad y ola.

**Graphics:** Pestaña en la que muestra un gráfico, seleccionado por el usuario, ampliado con su leyenda y una explicación de los valores de las abscisas, si procede.

**Help:** Pestaña en la que se muestra tanto el significado de las variables de las BB.DD en el caso de *View Costumers* y *View Analysis*, como unas pequeñas instrucciones de uso en el caso de las pestañas *Analysis Old Survey*, *Analysis New Survey*, *Independece Analysis* y *Waves Comparison*.

### 1.3.2.1 View Costumers

View Customers | View Products | Analysis Old Survey | Analysis New Survey | Independence Analysis | Waves Comparison | Graphics | Help

| OLA | CODENCU | IDCENTRO | ENSEÑA | ENCUESTA | DIA         | HORA | P1 | P2A | P2B | P |
|-----|---------|----------|--------|----------|-------------|------|----|-----|-----|---|
| 14  | 1007B   | 1007     | 1      | B        | 13227321600 | 0    | 1  | 1   | 1   | 1 |
| 14  | 2006A   | 2006     | 2      | A        | 13227321600 | 0    | 1  | 1   | 1   | 1 |
| 14  | 2009B   | 2009     | 2      | B        | 13227321600 | 0    | 1  | 1   | 1   | 1 |
| 14  | 4002A   | 4002     | 4      | A        | 13227321600 | 0    | 0  | 0   | 1   | 1 |
| 14  | 5001A   | 5001     | 5      | A        | 13227321600 | 0    | 1  | 1   | 0   | 0 |
| 14  | 7008B   | 7008     | 7      | B        | 13227321600 | 0    | 1  | 1   | 1   | 1 |
| 14  | 7012B   | 7012     | 7      | B        | 13227321600 | 0    | 1  | 1   | 1   | 1 |
| 14  | 1003B   | 1003     | 1      | B        | 13227321600 | 0    | 1  | 1   | 1   | 1 |
| 14  | 1007A   | 1007     | 1      | A        | 13227321600 | 0    | 1  | 1   | 0   | 0 |
| 14  | 4003A   | 4003     | 4      | A        | 13227321600 | 0    | 0  | 0   | 1   | 0 |
| 14  | 5002B   | 5002     | 5      | B        | 13227321600 | 0    | 1  | 1   | 0   | 1 |
| 14  | 9028A   | 9028     | 9      | A        | 13227321600 | 0    | 1  | 1   | 1   | 1 |
| 14  | 9031B   | 9031     | 9      | B        | 13227321600 | 0    | 1  | 1   | 0   | 1 |
| 14  | 2007B   | 2007     | 2      | B        | 13227321600 | 0    | 1  | 1   | 0   | 1 |
| 14  | 6007A   | 6007     | 6      | A        | 13227321600 | 0    | 1  | 0   | 0   | 1 |
| 14  | 7004B   | 7004     | 7      | B        | 13227321600 | 0    | 1  | 1   | 1   | 1 |
| 14  | 1004A   | 1004     | 1      | A        | 13227321600 | 0    | 1  | 0   | 0   | 0 |
| 14  | 1005A   | 1005     | 1      | A        | 13227321600 | 0    | 1  | 1   | 1   | 1 |

Olas que se puede ver en la encuesta antigua

Muestra la encuesta seleccionada

Haga clic en la pestaña View Products. para ver la siguiente pestaña

VIEW

Selección encuesta antigua para visualizar

Old Survey New Survey

Selección encuesta nueva para visualizar

Muestra en la pestaña Help el significado de las variables de la BB DD

HELP

### 1.3.2.2 View Products

View Customers | View Products | Analysis Old Survey | Analysis New Survey | Independence Analysis | Waves Comparison | Graphics | Help

| OLA | CENTRO | IDCENTRO | IDENSEÑA | IDPROD | MDMF | MARCA            | MARCA            | FABRICANTE | TTU | f |
|-----|--------|----------|----------|--------|------|------------------|------------------|------------|-----|---|
| 16  | 1000   | 1000     | 1        | 6      | MD   | EL MEJOR PRE...  | ALCAMPO.S.A. ... | 125        | 125 | 4 |
| 16  | 1000   | 1000     | 1        | 8      | MD   | EL MEJOR PRE...  | ALCAMPO.S.A. ... | 125        | 125 | 4 |
| 16  | 1000   | 1000     | 1        | 14     | MD   | EL MEJOR PRE...  | ALCAMPO.S.A. ... | 100        | 100 | 4 |
| 16  | 1000   | 1000     | 1        | 6      | MD   | AUCHAN           | ALCAMPO.S.A. ... | 125        | 125 | 4 |
| 16  | 1000   | 1000     | 1        | 11     | MD   | AUCHAN           | ALCAMPO.S.A. ... | 60         | 60  | 6 |
| 16  | 1000   | 1000     | 1        | 8      | MD   | AUCHAN           | ALCAMPO.S.A. ... | 125        | 125 | 4 |
| 16  | 1000   | 1000     | 1        | 101    | MD   | MEJOR PRECIO ... | 8480029          | 85         | 92  | 1 |
| 16  | 1000   | 1000     | 1        | 7      | MD   | RIK Y ROK        | ALCAMPO.S.A. ... | 100        | 100 | 4 |
| 16  | 1000   | 1000     | 1        | 14     | MF   | AUCHAN           | ALCAMPO.S.A. ... | 125        | 125 | 4 |
| 16  | 1000   | 1000     | 1        | 11     | MF   | REINA            | POSTRES Y DU...  | 100        | 100 | 4 |
| 16  | 1000   | 1000     | 1        | 11     | MF   | CLESA            | CLESA S.A. ...   | 60         | 55  | 6 |
| 16  | 1000   | 1000     | 1        | 11     | MF   | ...              | ...              | 60         | 60  | 6 |
| 16  | 1000   | 1000     | 1        | 91     | MD   | AUCHAN           | 8480029          | 1          | 1   | 1 |
| 16  | 1000   | 1000     | 1        | 14     | MF   | CLESA            | CLESA S.A. ...   | 100        | 105 | 4 |
| 16  | 1000   | 1000     | 1        | 6      | MF   | C.L.ASTURIANA... | C.A. PEÑASANT... | 125        | 125 | 4 |
| 16  | 1000   | 1000     | 1        | 8      | MF   | C.L.ASTURIANA... | C.A. PEÑASANT... | 125        | 125 | 4 |
| 16  | 1000   | 1000     | 1        | 106    | MF   | FRISCOS          | FRISCOS SA       | 92         | 92  | 3 |
| 16  | 1000   | 1000     | 1        | 11     | MF   | Danone           | DANONE ESPA...   | 60         | 60  | 6 |

Muestra la Ola seleccionada

Haga clic en la pestaña Analysis Old Survey. para ver la siguiente pestaña

VIEW

Selección de Ola que se quiere mostrar

WAVE 14 WAVE 15 WAVE 16 WAVE 17 WAVE 18 WAVE 19 WAVE 20

Muestra en la pestaña Help el significado de las variables de la BB DD

HELP

### 1.3.2.3 Analysis Old Survey

Analysis Old Survey

Store brands buys volume evolution

Study the evolution of consumer's buys in relation with store brands

Ejecuta el análisis seleccionado con la opción seleccionada

Execute

☒ Open the document

Mostrar el informe creado una vez acabada la ejecución del análisis

Help

Muestra los distintos análisis disponibles

Muestra en la pestaña help las instrucciones de uso de esta pestaña

The analysis has finished, You could continue making analysis

Enseña Age Gender Social status

Opciones disponibles para el análisis

Gráfico con los resultados del análisis

Haga clic en la pestaña Analysis New Survey. para ver la siguiente pestaña

### 1.3.2.4 Analysis New Survey

Analysis New Survey

Store brands characteristics: punctuation by supermarket format

Study about the punctuation of the characteristics on store brands by supermarket format

Ejecuta el análisis seleccionado con la opción seleccionada

Execute

☒ Open the document

Mostrar el informe creado una vez acabada la ejecución del análisis

Help

Muestra los distintos análisis disponibles

Muestra en la pestaña help las instrucciones de uso de esta pestaña

The document has been created. The analysis has finished, You could continue making analysis

Brand Age Gender Status Frequency visit

Opciones disponibles para el análisis

Gráfico con los resultados del análisis

Haga clic en la pestaña Independence Analysis. para ver la siguiente pestaña



### 1.3.2.5 Independence Analysis

View Customers | View Products | Analysis Old Survey | Analysis New Survey | Independence Analysis | Waves Comparison | Graphics | Help

Variable 1: KNOWLEDGE  
Variable 2: PUNCTUATION PACKAGING

Seleccionar Variable 1  
Seleccionar Variable 2

Survey: New Survey  
Execute

Open the document ☒

The independence analysis has finished, You could continue making independence analysis

Opciones disponibles para el alpha  
☐  $\alpha = 0.075$ 
☒  $\alpha = 0.05$ 
☐  $\alpha = 0.02$ 
☐  $\alpha = 0.01$

$\chi^2$

Haga clic en la pestaña Waves Comparison, para ver la siguiente pestaña

### 1.3.2.6 Waves Comparison

View Customers | View Products | Analysis Old Survey | Analysis New Survey | Independence Analysis | Waves Comparison | Graphics | Help

Decision analysis

WAVE 14  
WAVE 15  
WAVE 16  
WAVE 17  
WAVE 18  
WAVE 19

Knowledge: 0.4  
Buy/Knowledge: 0.2  
% People buying more: 0.2  
% People buying the same: 0.2

Please, select the weight of every Criteria  
Execute

Matrices intermedias del análisis

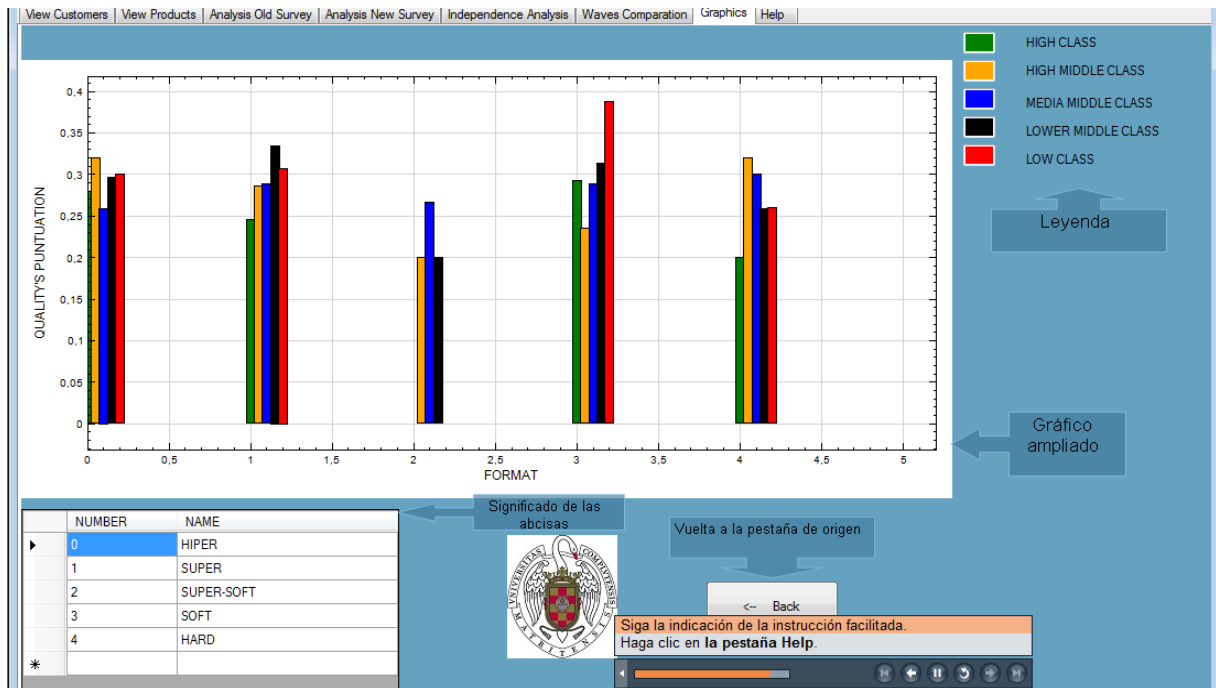
| Knowledge   | % Buy / knowledge  | Buy More   | Buy the same   | range of age                             | kind of matrix              |
|---|--|--|--|--|-----------------------------|
| (0.7 0.78 0.85)<br>(0.67 0.75 0.82)<br>(0.69 0.77 0.84) | (0.52 0.58 0.64)<br>(0.55 0.62 0.69)<br>(0.58 0.64 0.71) | (0.3 0.33 0.36)<br>(0.29 0.32 0.36)<br>(0.19 0.22 0.24)  | (0.39 0.44 0.48)<br>(0.47 0.52 0.57)<br>(0.53 0.59 0.65) | under 25<br>between 25 and 50<br>over 50 | performance matrix          |
| (0.28 0.31 0.34)<br>(0.27 0.3 0.33)<br>(0.28 0.31 0.34) | (0.1 0.12 0.13)<br>(0.11 0.12 0.14)<br>(0.12 0.13 0.14)  | (0.06 0.07 0.07)<br>(0.06 0.06 0.07)<br>(0.04 0.04 0.05) | (0.08 0.09 0.1)<br>(0.09 0.1 0.11)<br>(0.11 0.12 0.13)   | under 25<br>between 25 and 50<br>over 50 | weighted performance matrix |
| 0.69 / 0.31<br>0.7 / 0.3<br>0.69 / 0.31                 | 0.88 / 0.12<br>0.87 / 0.13                               | 0.93 / 0.07<br>0.94 / 0.06<br>0.96 / 0.04                | 0.91 / 0.09<br>0.9 / 0.1<br>0.88 / 0.12                  | under 25<br>between 25 and 50<br>over 50 | d* / di- matrix             |

Resultado del análisis

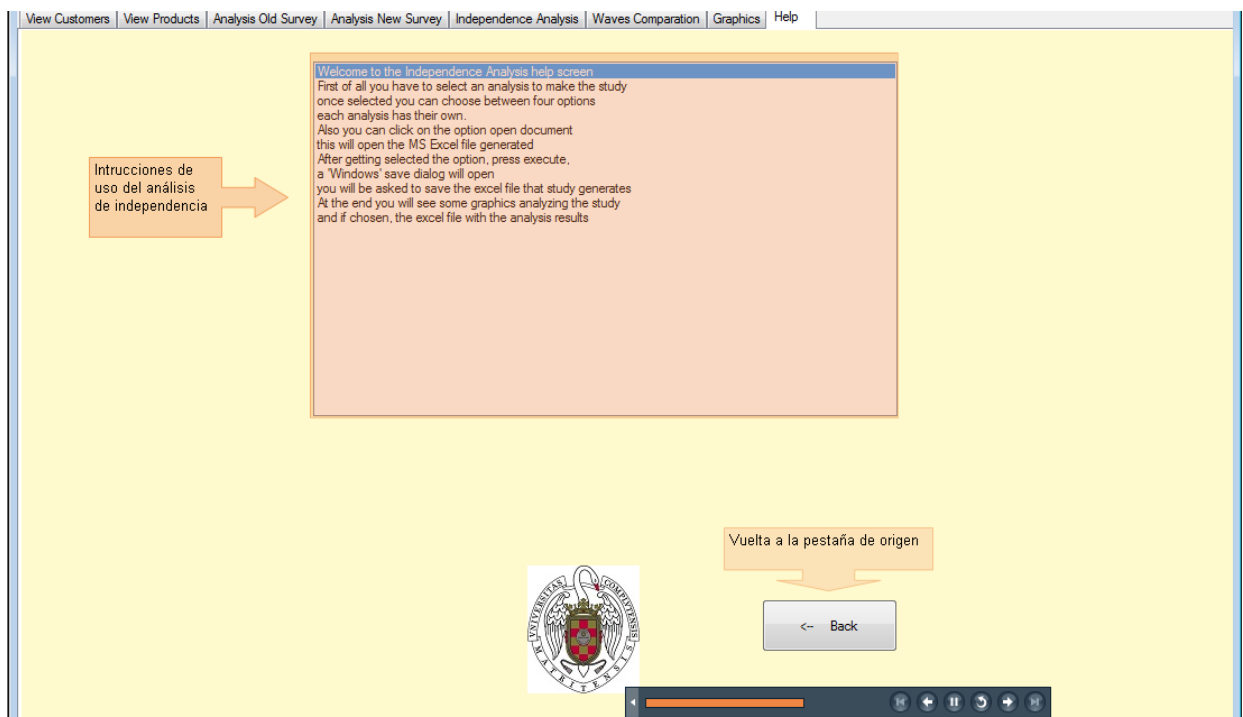
Best:  $P_i = 0.149$  ---> Over50 Top Consumer Age  
Middle:  $P_i = 0.148$  ---> Between 25 and 50 Middle consumer Age  
Worst:  $P_i = 0.145$  ---> Under 25 Less Consumer Age

Haga clic en la pestaña Graphics.

### 1.3.2.7 Graphics



### 1.3.2.8 Help

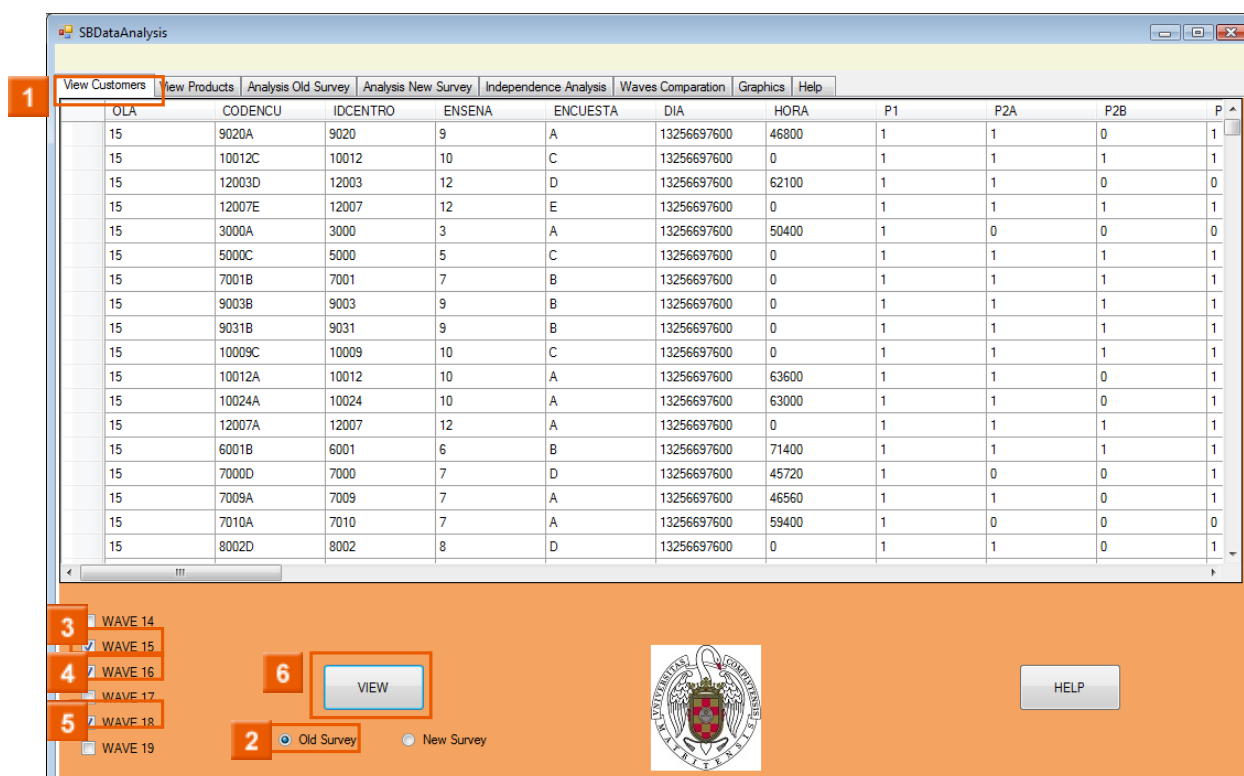


## 2 Visualización de las Bases de Datos

### 2.1 Visualización de BB.DD encuesta consumidores

#### 2.1.1 Visualización encuesta antigua

A continuación se muestra como se realiza la visualización de las BB.DD de encuestas antiguas de consumidores. En este caso queremos ver las Olas 15,16 y 18.



The screenshot shows the SBDDataAnalysis application window. The 'View Customers' tab is selected in the menu bar. Below the menu bar, there is a list of survey waves (OLA) with columns for OLA, CODENCU, IDCENTRO, ENSENA, ENCUESTA, DIA, HORA, P1, P2A, P2B, and P. The list shows waves 15, 16, and 18. At the bottom of the window, there is a section with a 'VIEW' button and radio buttons for 'Old Survey' and 'New Survey'. The 'Old Survey' radio button is selected. There are also checkboxes for 'WAVE 14', 'WAVE 15', 'WAVE 16', 'WAVE 17', 'WAVE 18', and 'WAVE 19'. The 'WAVE 15', 'WAVE 16', and 'WAVE 18' checkboxes are checked. A 'HELP' button is also visible.

1 Haga clic en la **pestaña View Customers**.

2 Haga clic en el **icono Old Survey**.

3 Haga clic en la **casilla WAVE 15**.

4 Haga clic en la **casilla WAVE 16**.

5 Haga clic en la **casilla WAVE 18**.

6 Haga clic en el **icono VIEW**. Se mostrará el tipo de encuestas y olas seleccionadas.

SBDDataAnalysis (No response)

View Customers

View Products

Analysis Old Survey

Analysis New Survey

Independence Analysis

Waves Comparison

Graphics

Help

|  | OLA | CODENCU | IDCENTRO | ENSENA | ENCUESTA | DIA         | HORA  | P1 | P2A | P2B | P |
|--|-----|---------|----------|--------|----------|-------------|-------|----|-----|-----|---|
|  | 15  | 8000E   | 8000     | 8      | E        | 13256697600 | 48000 | 0  | 0   | 0   | 0 |
|  | 15  | 8003A   | 8003     | 8      | A        | 13256697600 | 70200 | 1  | 1   | 0   | 1 |
|  | 15  | 9014B   | 9014     | 9      | B        | 13256697600 | 0     | 1  | 1   | 0   | 1 |
|  | 15  | 9307D   | 9307     | 9      | D        | 13256697600 | 0     | 1  | 1   | 1   | 1 |
|  | 15  | 2002D   | 2002     | 2      | D        | 13256697600 | 39600 | 1  | 1   | 0   | 1 |
|  | 15  | 2002I   | 2002     | 2      | I        | 13256697600 | 39600 | 1  | 1   | 0   | 1 |
|  | 15  | 2004C   | 2004     | 2      | C        | 13256697600 | 0     | 1  | 1   | 0   | 1 |
|  | 15  | 10005C  | 10005    | 10     | C        | 13256697600 | 0     | 1  | 1   | 1   | 1 |
|  | 15  | 10009E  | 10009    | 10     | E        | 13256697600 | 0     | 1  | 1   | 1   | 0 |
|  | 15  | 10025B  | 10025    | 10     | B        | 13256697600 | 0     | 1  | 1   | 1   | 1 |
|  | 15  | 10025D  | 10025    | 10     | D        | 13256697600 | 0     | 1  | 1   | 1   | 1 |
|  | 15  | 12006B  | 12006    | 12     | B        | 13256697600 | 0     | 1  | 1   | 0   | 1 |
|  | 15  | 12007D  | 12007    | 12     | D        | 13256697600 | 0     | 1  | 1   | 0   | 1 |
|  | 15  | 12008A  | 12008    | 12     | A        | 13256697600 | 66600 | 1  | 1   | 0   | 0 |
|  | 15  | 2000B   | 2000     | 2      | B        | 13256697600 | 58200 | 1  | 1   | 1   | 1 |
|  | 15  | 2001B   | 2001     | 2      | B        | 13256697600 | 0     | 1  | 1   | 0   | 1 |
|  | 15  | 2015B   | 2015     | 2      | B        | 13256697600 | 73200 | 1  | 1   | 1   | 1 |
|  | 15  | 4005C   | 4005     | 4      | C        | 13256697600 | 0     | 1  | 0   | 0   | 1 |

WAVE 14

WAVE 15

WAVE 16

WAVE 17


WAVE 18

WAVE 19

VIEW

Old Survey

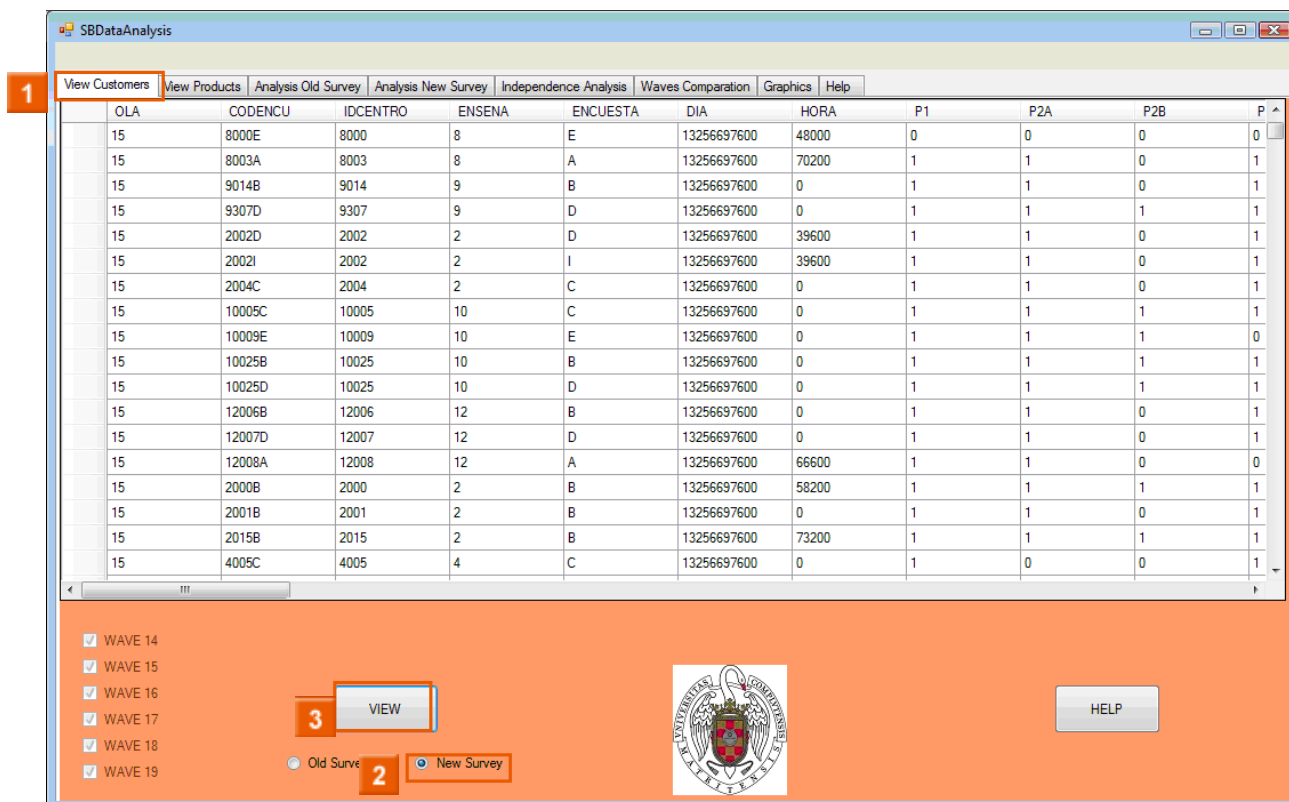
New Survey



HELP

## 2.1.2 Visualización encuesta nueva

A continuación se muestra como se realiza la visualización de las BB.DD de encuestas nuevas de consumidores.



The screenshot shows the SBDDataAnalysis application window. The 'View Customers' tab is selected, indicated by a red box and the number 1. Below the tab bar is a table with the following columns: OLA, CODENCU, IDCENTRO, ENSENA, ENCUESTA, DIA, HORA, P1, P2A, P2B, and P. The table contains 20 rows of data. At the bottom of the window, there is a section with a list of waves (WAVE 14 to WAVE 19) with checkboxes, a 'VIEW' button (highlighted with a red box and the number 3), a 'New Survey' radio button (highlighted with a red box and the number 2), and a 'HELP' button. A logo is also present in this section.

| OLA | CODENCU | IDCENTRO | ENSENA | ENCUESTA | DIA         | HORA  | P1 | P2A | P2B | P |
|-----|---------|----------|--------|----------|-------------|-------|----|-----|-----|---|
| 15  | 8000E   | 8000     | 8      | E        | 13256697600 | 48000 | 0  | 0   | 0   | 0 |
| 15  | 8003A   | 8003     | 8      | A        | 13256697600 | 70200 | 1  | 1   | 0   | 1 |
| 15  | 9014B   | 9014     | 9      | B        | 13256697600 | 0     | 1  | 1   | 0   | 1 |
| 15  | 9307D   | 9307     | 9      | D        | 13256697600 | 0     | 1  | 1   | 1   | 1 |
| 15  | 2002D   | 2002     | 2      | D        | 13256697600 | 39600 | 1  | 1   | 0   | 1 |
| 15  | 2002I   | 2002     | 2      | I        | 13256697600 | 39600 | 1  | 1   | 0   | 1 |
| 15  | 2004C   | 2004     | 2      | C        | 13256697600 | 0     | 1  | 1   | 0   | 1 |
| 15  | 10005C  | 10005    | 10     | C        | 13256697600 | 0     | 1  | 1   | 1   | 1 |
| 15  | 10009E  | 10009    | 10     | E        | 13256697600 | 0     | 1  | 1   | 1   | 0 |
| 15  | 10025B  | 10025    | 10     | B        | 13256697600 | 0     | 1  | 1   | 1   | 1 |
| 15  | 10025D  | 10025    | 10     | D        | 13256697600 | 0     | 1  | 1   | 1   | 1 |
| 15  | 12006B  | 12006    | 12     | B        | 13256697600 | 0     | 1  | 1   | 0   | 1 |
| 15  | 12007D  | 12007    | 12     | D        | 13256697600 | 0     | 1  | 1   | 0   | 1 |
| 15  | 12008A  | 12008    | 12     | A        | 13256697600 | 66600 | 1  | 1   | 0   | 0 |
| 15  | 2000B   | 2000     | 2      | B        | 13256697600 | 58200 | 1  | 1   | 1   | 1 |
| 15  | 2001B   | 2001     | 2      | B        | 13256697600 | 0     | 1  | 1   | 0   | 1 |
| 15  | 2015B   | 2015     | 2      | B        | 13256697600 | 73200 | 1  | 1   | 1   | 1 |
| 15  | 4005C   | 4005     | 4      | C        | 13256697600 | 0     | 1  | 0   | 0   | 1 |

1 Haga clic en la **pestaña View Customers**.

2 Haga clic en el **icono New Survey**.

3 Haga clic en el **icono VIEW**. Se mostrará el tipo de encuesta seleccionada.

SBDDataAnalysis

View Customers

View Products

Analysis Old Survey

Analysis New Survey

Independence Analysis

Waves Comparison

Graphics

Help

|  | OLA | CODENCU | IDCENTRO | ENSENA | FORMATO | ENCUESTA | DIA        | HORA    | P1   | P2A | P |
|--|-----|---------|----------|--------|---------|----------|------------|---------|------|-----|---|
|  | 20  | 2015B   | 2015     | 4      | 1       | B        | 39124      | 61200   | 1    | 0   | 1 |
|  | 20  | 6027B   | 6027     | 12     | 4       | B        | 39124      | 59400   | 1    | 1   | 0 |
|  | 20  | 2010C   | 2010     | 4      | 1       | C        | 30/11/2007 | 46800   | 1    | 1   | 1 |
|  | 20  | 2011I   | 2011     | 4      | 1       | I        | 39152      | 54780   | 1    | 1   | 1 |
|  | 20  | 4004B   | 4004     | 6      | 2       | B        | 30/11/2007 | 63000   | 1    | 1   | 0 |
|  | 20  | 9049A   | 9049     | 5      | 4       | A        | 30/11/2007 | 64800   | 1    | 1   | 1 |
|  | 20  | 9141A   | 9141     | 5      | 4       | A        | 30/11/2007 | 68400   | 1    | 1   | 1 |
|  | 20  | 9201B   | 9201     | 5      | 4       | B        |            | null    | 1    | 1   | 1 |
|  | 20  | 9225H   | 9225     | 5      | 4       | H        | 39213      | 1257120 | 1    | 1   | 1 |
|  | 20  | 10024C  | 10024    | 13     | 2       | C        |            | null    | 1    | 0   | 0 |
|  | 20  | 14001B  | 14001    | 10     | 5       | B        | 30/11/2007 | 1308960 | 0    | 1   | 1 |
|  | 20  | 14001C  | 14001    | 10     | 5       | C        | 30/11/2007 | 1296000 | 1    | 1   | 0 |
|  | 20  | 14012A  | 14012    | 10     | 5       | A        | 39183      | 46800   | null | 1   | 0 |
|  | 20  | 14020C  | 14020    | 10     | 5       | C        | 39213      | 70200   | 1    | 1   | 1 |
|  | 20  | 15104B  | 15104    | 1      | 2       | B        | 29/11/2007 | 64800   | 1    | 1   | 0 |
|  | 20  | 1000A   | 1000     | 2      | 1       | A        | 30/11/2007 | 63900   | 1    | 1   | 0 |
|  | 20  | 1001B   | 1001     | 2      | 1       | B        |            | null    | 1    | 1   | 0 |
|  | 20  | 1002D   | 1002     | 2      | 1       | D        | 39416      | 49380   | 1    | 1   | 1 |

WAVE 14

WAVE 15

WAVE 16

WAVE 17


WAVE 18

WAVE 19

VIEW

Old Survey

New Survey



HELP

### 2.1.3 Ver significado de las variables de consumidores

A continuación se muestra cómo ver el significado de las variables de las BB.DD de encuestas a consumidores, tanto de la antigua encuesta como de la nueva.

| OLA | CODENCU | IDCENTRO | ENSENA | FORMATO | ENCUESTA | DIA        | HORA    | P1   | P2A | P |
|-----|---------|----------|--------|---------|----------|------------|---------|------|-----|---|
| 20  | 2015B   | 2015     | 4      | 1       | B        | 39124      | 61200   | 1    | 0   | 1 |
| 20  | 6027B   | 6027     | 12     | 4       | B        | 39124      | 59400   | 1    | 1   | 0 |
| 20  | 2010C   | 2010     | 4      | 1       | C        | 30/11/2007 | 46800   | 1    | 1   | 1 |
| 20  | 2011I   | 2011     | 4      | 1       | I        | 39152      | 54780   | 1    | 1   | 1 |
| 20  | 4004B   | 4004     | 6      | 2       | B        | 30/11/2007 | 63000   | 1    | 1   | 0 |
| 20  | 9049A   | 9049     | 5      | 4       | A        | 30/11/2007 | 64800   | 1    | 1   | 1 |
| 20  | 9141A   | 9141     | 5      | 4       | A        | 30/11/2007 | 68400   | 1    | 1   | 1 |
| 20  | 9201B   | 9201     | 5      | 4       | B        |            | null    | 1    | 1   | 1 |
| 20  | 9225H   | 9225     | 5      | 4       | H        | 39213      | 1257120 | 1    | 1   | 1 |
| 20  | 10024C  | 10024    | 13     | 2       | C        |            | null    | 1    | 0   | 0 |
| 20  | 14001B  | 14001    | 10     | 5       | B        | 30/11/2007 | 1308960 | 0    | 1   | 1 |
| 20  | 14001C  | 14001    | 10     | 5       | C        | 30/11/2007 | 1296000 | 1    | 1   | 0 |
| 20  | 14012A  | 14012    | 10     | 5       | A        | 39183      | 46800   | null | 1   | 0 |
| 20  | 14020C  | 14020    | 10     | 5       | C        | 39213      | 70200   | 1    | 1   | 1 |
| 20  | 15104B  | 15104    | 1      | 2       | B        | 29/11/2007 | 64800   | 1    | 1   | 0 |
| 20  | 1000A   | 1000     | 2      | 1       | A        | 30/11/2007 | 63900   | 1    | 1   | 0 |
| 20  | 1001B   | 1001     | 2      | 1       | B        |            | null    | 1    | 1   | 0 |
| 20  | 1002D   | 1002     | 2      | 1       | D        | 39416      | 49380   | 1    | 1   | 1 |

**1** Haga clic en el icono **HELP**. Se mostrará la pestaña *Help* con el significado de las variables de las BB.DD.

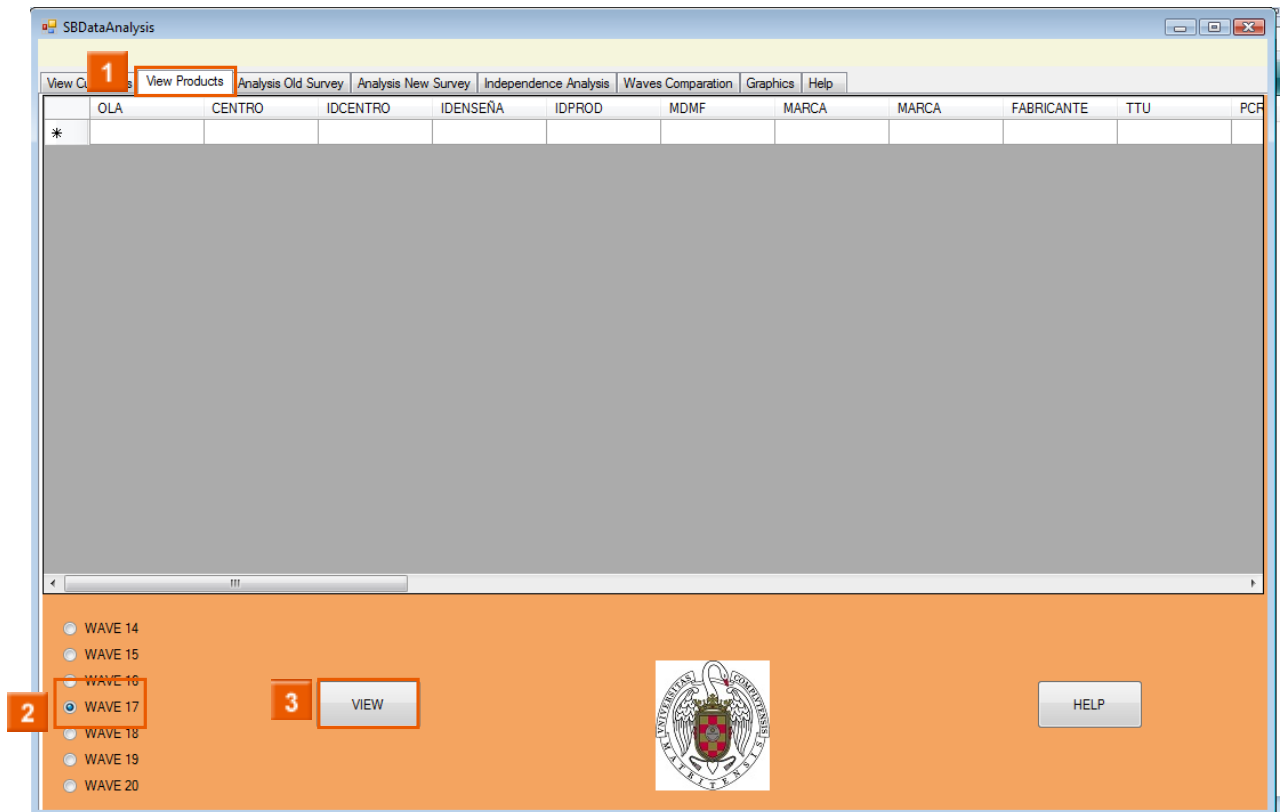
| Variable | Significado  |
|----------|--|
| WAVE     | 14, 15, 16, 17, 18, 19   |
| CODENCU  | CENTRE'S NUMBER Y ENCUESTA   |
| IDCENTRO | CENTRE'S NUMBER  |
| ENSENA   | COMPANY'S NUMBER   |
| DIA      | DAY  |
| HORA     | HOUR   |
| FORMATO  | CENTRE'S FORMAT NUMBER *ONLY IN THE NEW SURVEY   |
| ENCUESTA | FROM A TO L  |
| P1       | P1 - ¿Do you know, although you have never bought it, food products which has the supermarket's store brand? |
| P2A      | P2 do you know a - Milk  |
| P2B      | P2 do you buy a - Milk   |
| P3A      | P3 do you know b - Oils  |
| P3B      | P3 do you buy b - Oils   |
| P4A      | P4 do you know c - Coffees   |
| P4B      | P4 do you buy c - Coffees  |
| P5A      | P5 do you know d - Chocolates  |
| P5B      | P5 do you buy d - Chocolates   |
| P6A      | P6 do you know e - Mermelades and honey  |
| P6B      | P6 do you buy e - Mermelades and honey   |
| P7A      | P7 do you know f - Biscuits and bread  |
| P7B      | P7 do you buy f - Biscuits and bread   |
| P8A      | P8 do you know g - Legumbres y varios  |
| P8B      | P8 do you buy g - Legumbres y varios   |
| P9A      | P9 do you know h - Conservas de pescado  |
| P9B      | P9 do you buy h - Conservas de pescado   |
| P10A     | P10 do you know i - Conservas vegetales  |
| P10B     | P10 do you buy i - Conservas vegetales   |
| P11A     | P11 do you know j - Pastas   |
| P11B     | P11 do you buy j - Pastas  |
| P12A     | P12 do you know k - Drinks   |
| P12B     | P12 do you buy k - Drinks  |
| P13A     | P13 do you know l - Turrones   |
| P13B     | P13 do you buy l - Turrones  |
| P14A     | P14 Which are your reasons to buy in this supermarket? Low prices  |

**1** Haga clic en el icono **Back** para volver a la pestaña *View Costumers*.

## 2.2 Visualización de BB.DD de productos de MDD

### 2.2.1 Visualización de una Ola de productos

A continuación se muestra cómo realizar la visualización de una ola de productos de MDD. Cada Ola se encuentra en una BB.DD distinta por lo que sólo se podrá visualizar una ola. En este caso queremos ver la Ola 17.



**1** Haga clic en la **pestaña View Products**.

**2** Haga clic en el **icono WAVE 17**.

**3** Haga clic en el **icono VIEW**. Se mostrará la ola de productos seleccionada.



- 125 -

## 2.2.2 Ver significado de las variables de los productos

A continuación se muestra cómo visualizar el significado de las diferentes variables de las BB.DD de los productos de MDD.

| OLA | CENTRO   | IDCENTRO | IDENSEÑA | IDPROD | MDMF | MARCA     | MARCA | FABRICANTE         | TTU |
|-----|----------|----------|----------|--------|------|-----------|-------|--------------------|-----|
| 17  | 9194EG15 | 9194     | 9        | 53     | MD   | DIA       | ...   | U.E. Para DIA S... | 0.2 |
| 17  | 2007J7   | 2007     | 2        | 53     | MF   | FEIRACO   | ...   | SOC.COOPERA...     | 0.2 |
| 17  | 12014EG8 | 12014    | 12       | 53     | MD   | CONSUMER  | ...   | RSI 1502965/S      | 0.2 |
| 17  | 2013EG8  | 2013     | 2        | 53     | MD   | CARREFOUR | ...   | RSI 1500345/SS...  | 0.2 |
| 17  | 7002EG8  | 7002     | 7        | 53     | MD   | CHAMPION  | ...   | RSI 1500345/S      | 0.2 |
| 17  | 1000EG4  | 1000     | 1        | 53     | MF   | PULEVA    | ...   | PULEVA FOODS...    | 0.2 |
| 17  | 1002M5   | 1002     | 1        | 53     | MF   | PULEVA    | ...   | PULEVA FOODS...    | 0.2 |
| 17  | 2005M7   | 2005     | 2        | 53     | MF   | PULEVA    | ...   | PULEVA FOODS...    | 0.2 |
| 17  | 2010M2   | 2010     | 2        | 53     | MF   | PULEVA    | ...   | PULEVA FOODS...    | 0.2 |
| 17  | 2013EG8  | 2013     | 2        | 53     | MF   | PULEVA    | ...   | PULEVA FOODS...    | 0.2 |
| 17  | 2018EG13 | 2018     | 2        | 53     | MF   | PULEVA    | ...   | PULEVA FOODS...    | 0.2 |
| 17  | 6001EG5  | 6001     | 6        | 53     | MF   | PULEVA    | ...   | PULEVA FOODS...    | 0.2 |
| 17  | 6006J2   | 6006     | 6        | 53     | MF   | PULEVA    | ...   | PULEVA FOODS...    | 0.2 |
| 17  | 6016EG9  | 6016     | 6        | 53     | MF   | PULEVA    | ...   | PULEVA FOODS...    | 0.2 |
| 17  | 6017EG9  | 6017     | 6        | 53     | MF   | PULEVA    | ...   | PULEVA FOODS...    | 0.2 |
| 17  | 6022MD9  | 6022     | 6        | 53     | MF   | PULEVA    | ...   | PULEVA FOODS...    | 0.2 |
| 17  | 6099EG15 | 6099     | 6        | 53     | MF   | PULEVA    | ...   | PULEVA FOODS...    | 0.2 |
| 17  | 9194EG15 | 9194     | 9        | 53     | MF   | PULEVA    | ...   | PULEVA FOODS...    | 0.2 |

WAVE 14  
WAVE 15  
WAVE 16  
**WAVE 17**  
WAVE 18  
WAVE 19  
WAVE 20

VIEW

1

HELP

**1** Haga clic en **Help**. Se mostrará la pestaña *Help* con el significado de las variables de las BB.DD.

| OLA                               | WAVE'S NUMBER: 14, 15, 16, 17, 18, 19  |
|-----------------------------------|--|
| CENTRO                            | CENTER'S NUMBER                        |
| IDCENTRO                          | CENTER'S NAME                          |
| IDENSEÑA                          | SURVEY'S ID                            |
| IDPROD                            | PRODUCT'S NAME                         |
| MDMF                              | STORE BRAND OR PRIVATE BRAND?          |
| MARCA                             | BRAND'S NAME                           |
| FABRICANTE                        | MANUFACTURER'S NAME                    |
| TTU                               | CAPACITY'S OR WEIGHT'S PRODUCT         |
| PCRUI                             | REAL CAPACITY'S OR WEIGHT'S PRODUCT    |
| NUR                               | NUMBER OF UNITS IN THE PRODUCT         |
| PRR                               | REAL PRICE                             |
| PRU                               | UNITS PRICE                            |
| LIDER                             | LEADER'S NAME                          |
| ML                                | BETTER PACKING THAN LEADER             |
| QL                                | SAME PACKING THAN LEADER               |
| PQL                               | WORST PACKING THAN LEADER              |
| ESPM13                            | SPACE BRAND >1/3                       |
| ESPM13                            | SPACE BRAND <1/3                       |
| ESPM13                            | SPACE BRAND =1/3                       |
| TECHO                             | HEIGHT'S PRODUCT IS IN THE TOP         |
| OJOS                              | HEIGHT'S PRODUCT IS IN THE EYE HEIGHT  |
| MANOS                             | HEIGHT'S PRODUCT IS IN THE HAND HEIGHT |
| SUELO                             | HEIGHT'S PRODUCT IS IN THE BOTTOM      |
| cabecera                          | PRODUCT IS IN THE 'CABECERA'           |
| ala                               | PRODUCT IS IN THE 'ALA'                |
| chimenea                          | PRODUCT IS IN THE 'CHIMENEA'           |
| otros                             | PRODUCT IS IN THE 'OTHER'              |
| Folleto                           | EXIST BROCHURE'S PRODUCT?              |
| PUBLICIDADCENTROVIDEOMEGAFONIASCT | EXIST COMMERCIAL'S PRODUCT?            |
| Carteles                          | EXIST POSTER'S PRODUCT?                |
| degustaciones                     | EXIST TASTE'S PRODUCT?                 |
| Otras                             | EXIST OTHER PROMOTION'S PRODUCT?       |
| SINPROMO                          | NOT EXIST PROMOTIONS                   |

1

Back

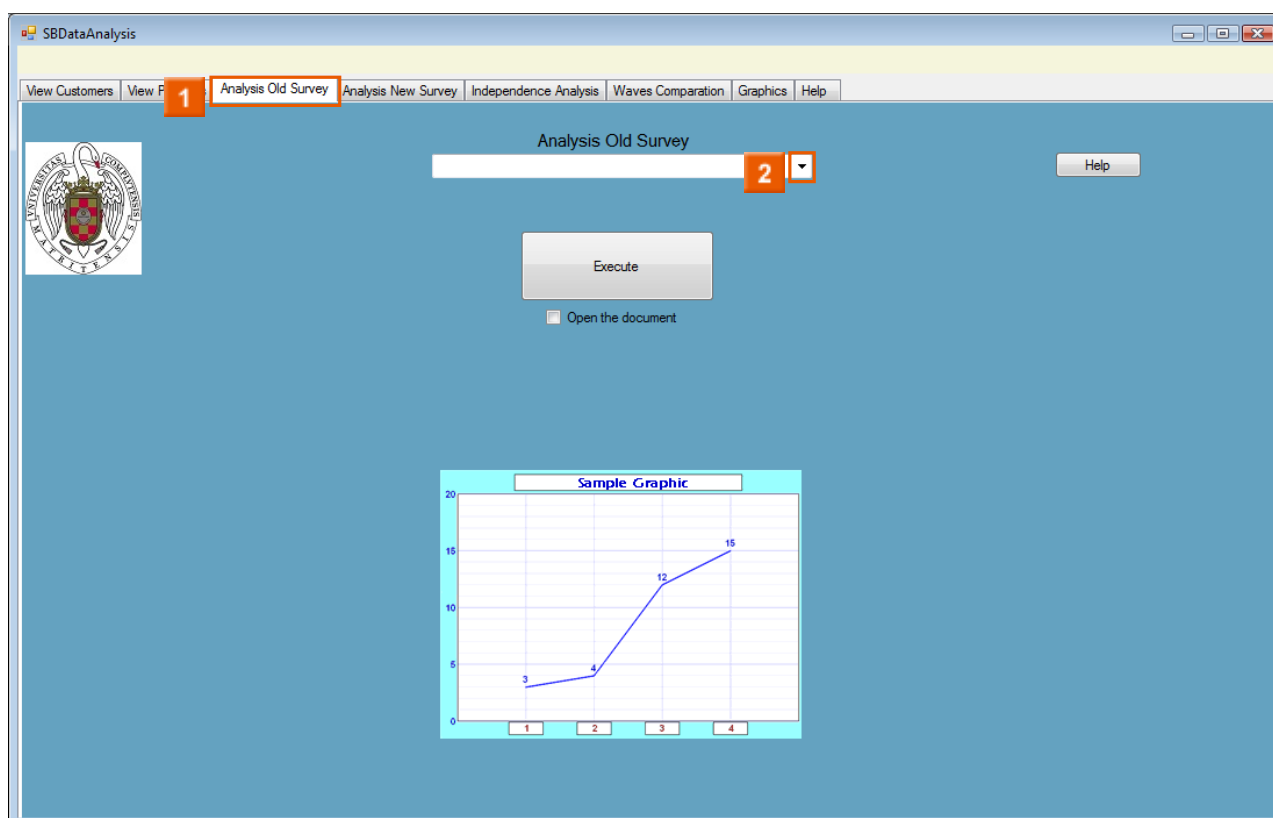
**1** Haga clic en el icono **Back** para volver a la pestaña *View Products*.

### 3 Análisis de encuestas de consumidores

#### 3.1 Análisis de encuesta de consumidores antigua

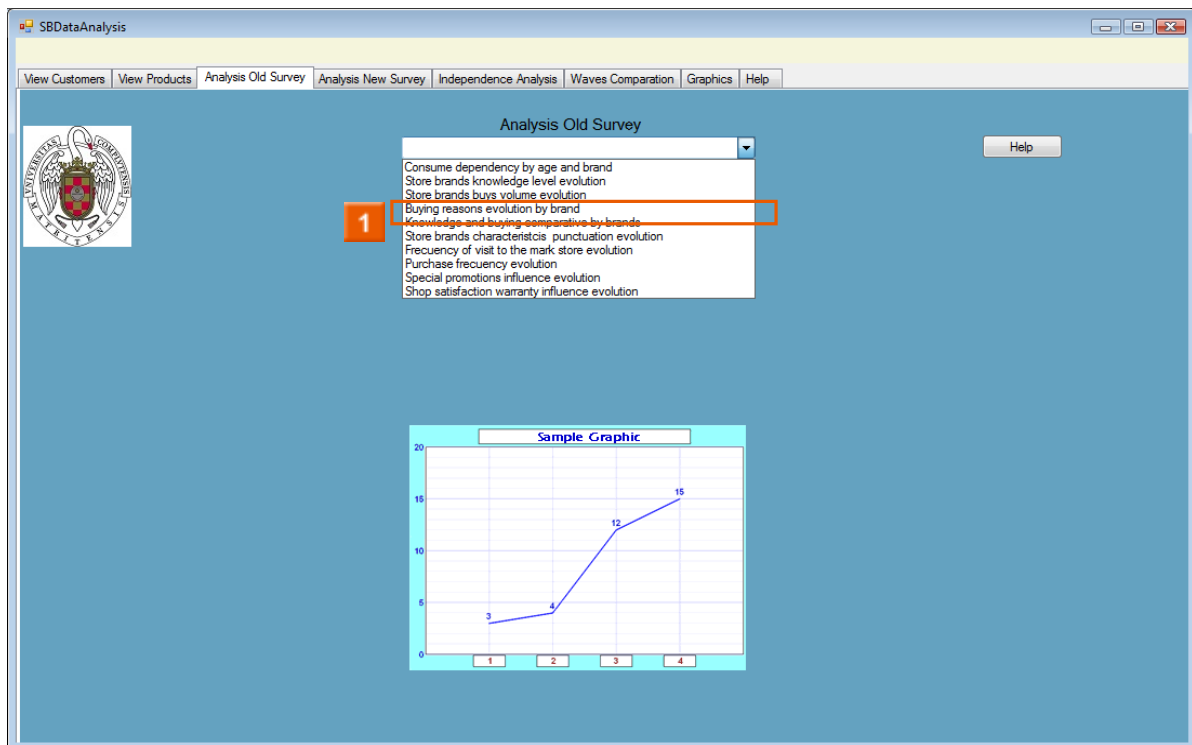
##### 3.1.1 Hacer un análisis de las encuestas antiguas a consumidores

A continuación se muestra cómo realizar un análisis de las encuestas antiguas de consumidores. Se creará un informe en formato Excel y se mostrará un gráfico con los resultados del análisis. En este ejemplo se realizará un análisis de la evolución de la compra de MDD por rangos de edad (menor de 25 años, entre 25 y 50 años, mayor de 50 años) de los consumidores. Además se quiere que al acabar el análisis, se abra el informe con los resultados del análisis.

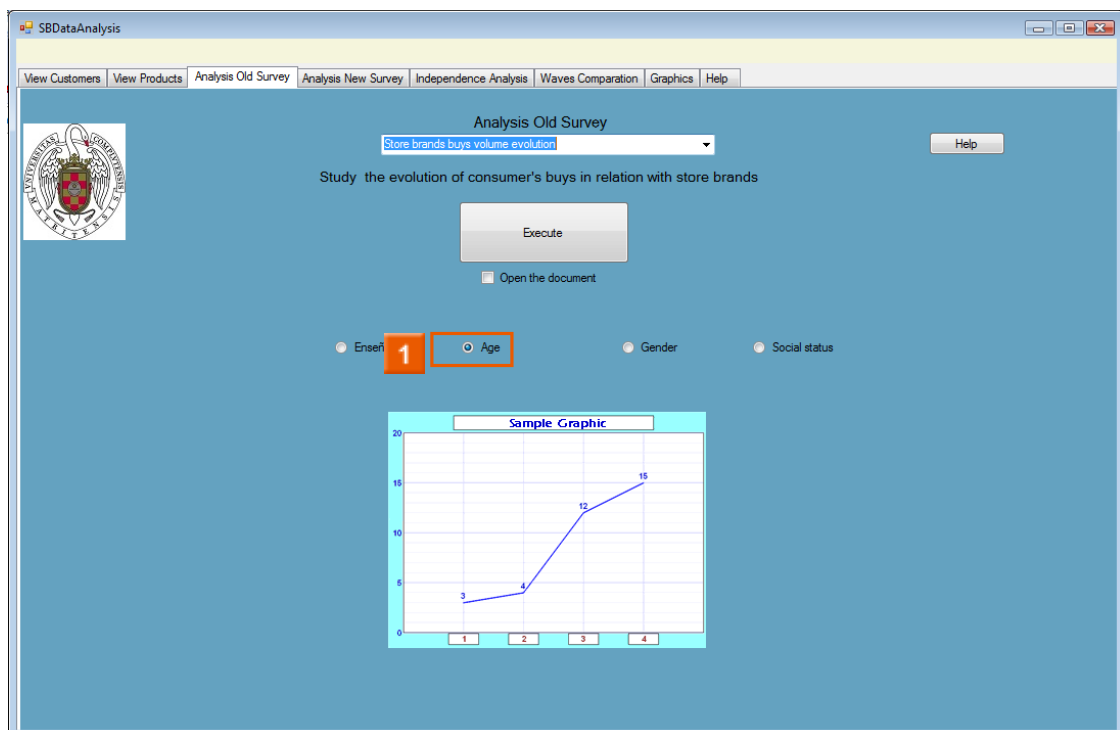


**1** Haga clic en la pestaña **Analysis Old Survey**.

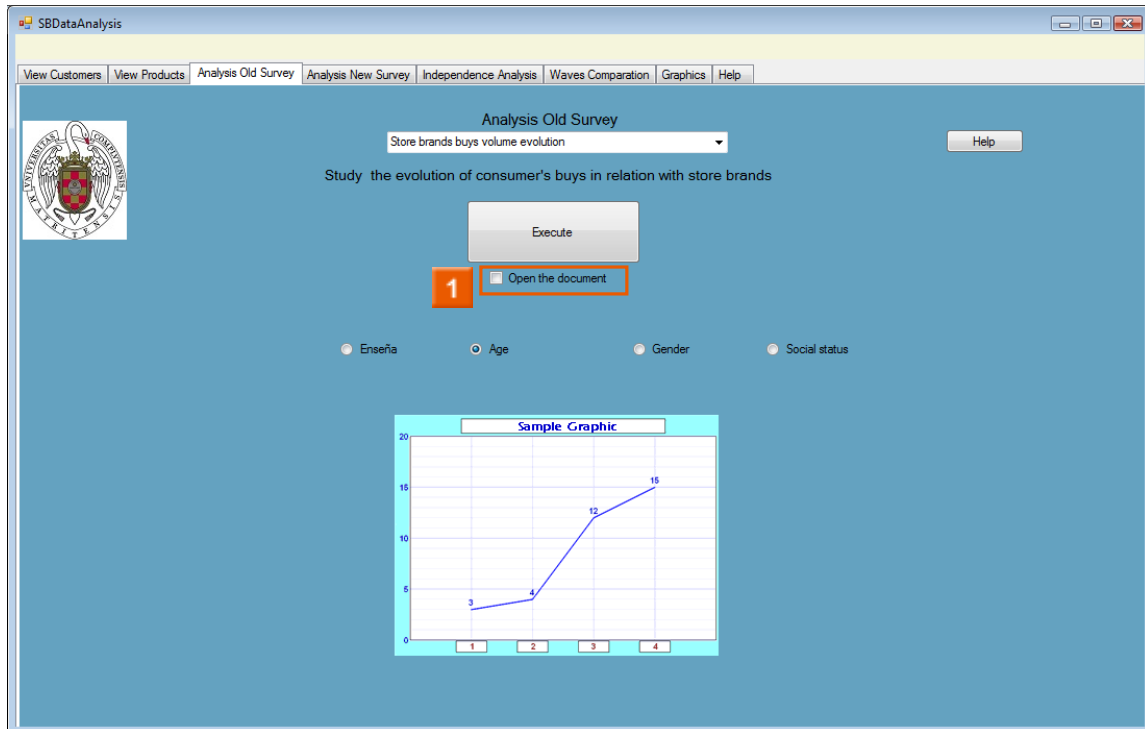
**2** Haga clic en ▼ para mostrar los diferentes estudios disponibles.



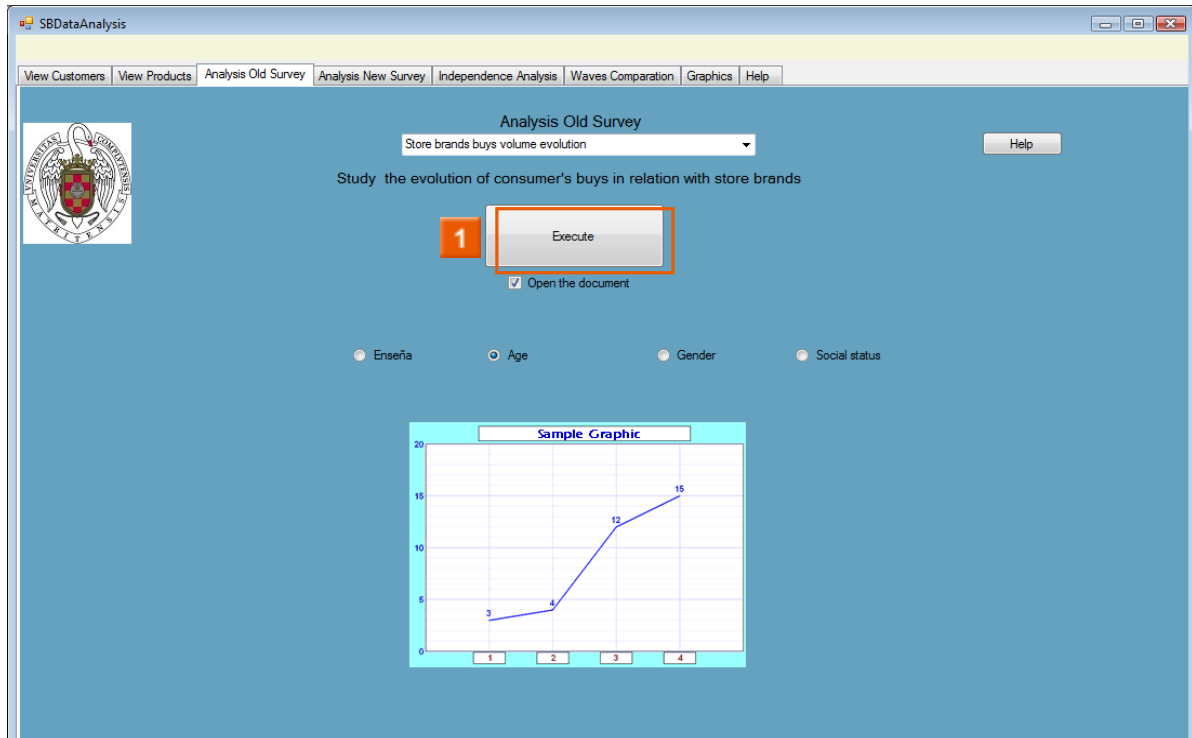
**1** Haga clic en el análisis **Store brands buys volume evolution**.



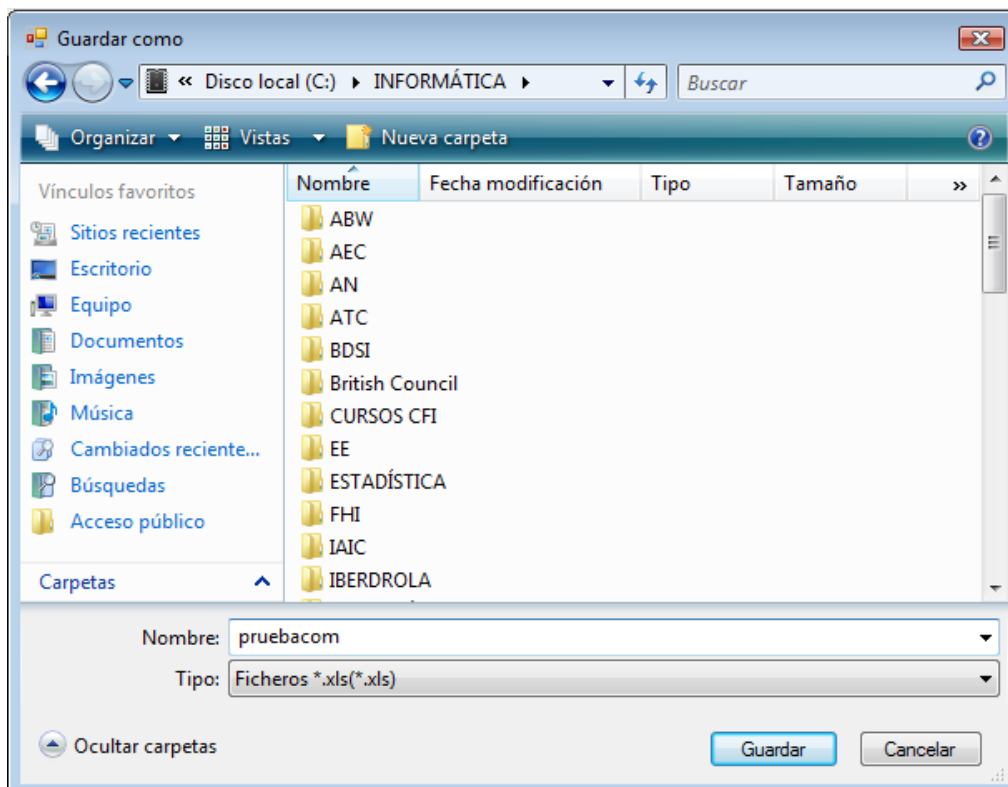
**1** Haga clic en **Age**.



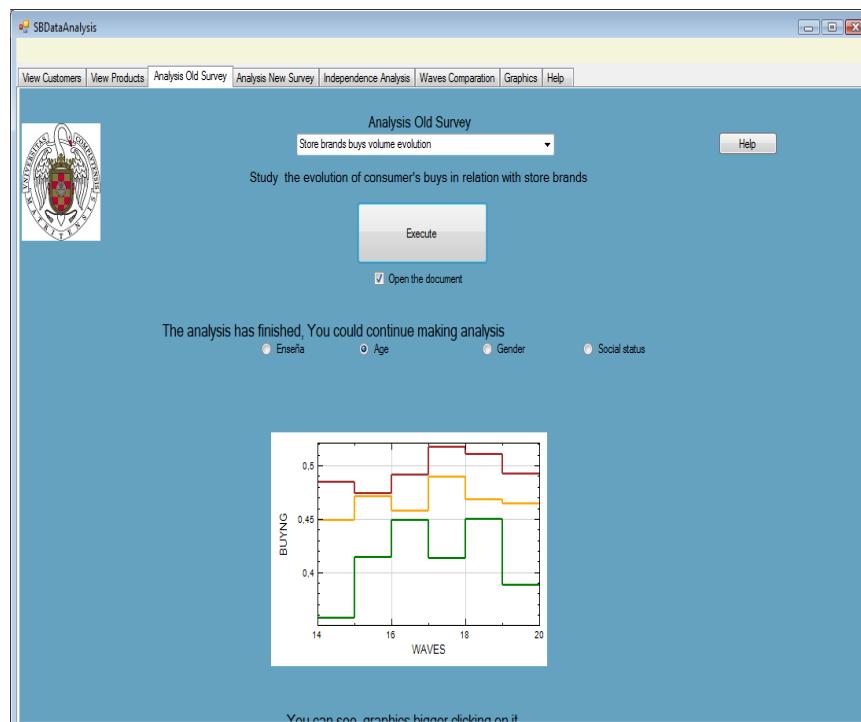
- 1 Haga clic en la casilla **Open the document** para que se abra el informe al acabar el estudio



- 1 Haga clic en el icono **Execute**.



Guarde el informe en la carpeta deseada. Al acabar el análisis, como se ha seleccionado la opción *Open the document* se abrirá el informe y se mostrará un gráfico con los resultados del análisis.



Microsoft Excel - prueba compra

Archivo Edición Ver Insertar Formato Herramientas Datos Ventana ?

Esciba una pregunta

10 Arial

16 Elapsed Time 0:00:00.125

17

18 Case Processing Summary

19

20 Cases

21 Included Excluded Total

22 COMPRAR \*OLA\*EDAD 3898 100.0% 0 .0% 3898 100.0%

23

24 Report

25 COMPRAR

26 OLA EDAD N Mean Harmonic Mean Geometric Mean Variance Std. Deviation Minimum Maximum

27 14 menor de 25 años 10 .3583<sup>a</sup> .0000 .059 24232 .00

28 entre 25 y 50 años 61 .4495<sup>a</sup> .0000 .059 24356 .00

29 mayor de 50 años 35 .4857<sup>a</sup> .0000 .058 24126 .00

30 Total 106 .4628<sup>a</sup> .0000 .059 24289 .00

31 15 menor de 25 años 57 .4152<sup>a</sup> .0000 .040 20013 .00

32 entre 25 y 50 años 230 .4725<sup>a</sup> .0000 .076 27517 .00

33 mayor de 50 años 134 .4751<sup>a</sup> .0000 .079 28135 .00

34 5 3 .4167<sup>a</sup> .3529 .049 22048 .25

35 Total 424 .4652<sup>a</sup> .0000 .072 26801 .00

36 16 menor de 25 años 94 .4495<sup>a</sup> .0000 .056 23640 .00

37 entre 25 y 50 años 337 .4595<sup>a</sup> .0000 .065 25476 .00

38 mayor de 50 años 194 .4923<sup>a</sup> .0000 .062 24880 .00

39 Total 625 .4676<sup>a</sup> .0000 .063 25941 .00

40 17 menor de 25 años 102 .4142<sup>a</sup> .0000 .045 21478 .00

41 entre 25 y 50 años 359 .4805<sup>a</sup> .0000 .071 26075 .00

42 mayor de 50 años 177 .5194<sup>a</sup> .0000 .065 25550 .00

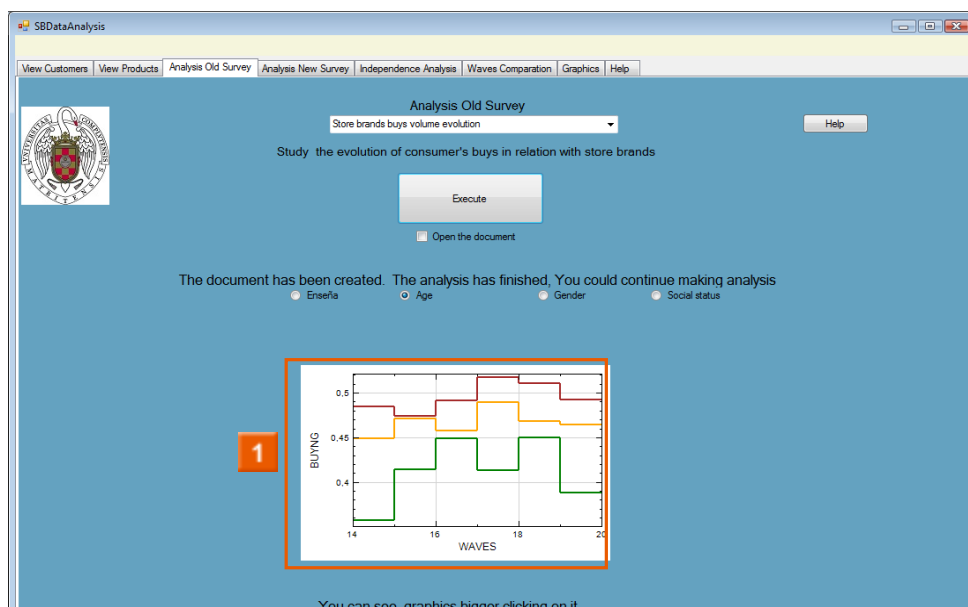
43 5 4 .5208<sup>a</sup> .0000 .178 42150 .00

Sheet1

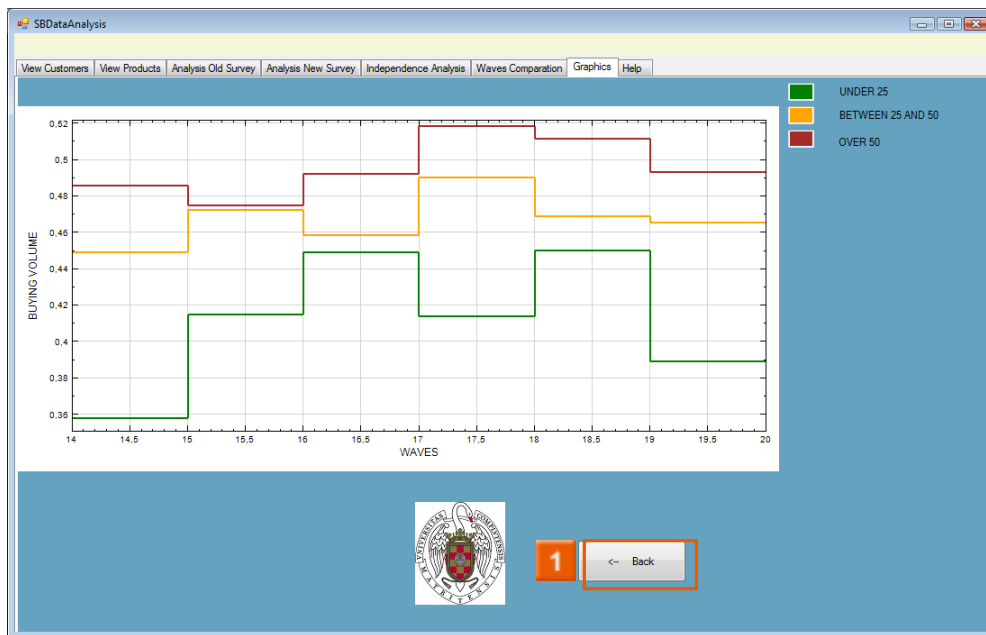
Listo

### 3.1.2 Ampliar gráfico de resultado

A continuación se muestra cómo ampliar un gráfico de resultados de un análisis. Con esta ampliación el gráfico se verá más grande y se mostrará la leyenda y una tabla con el significado de las abscisas si es necesario.



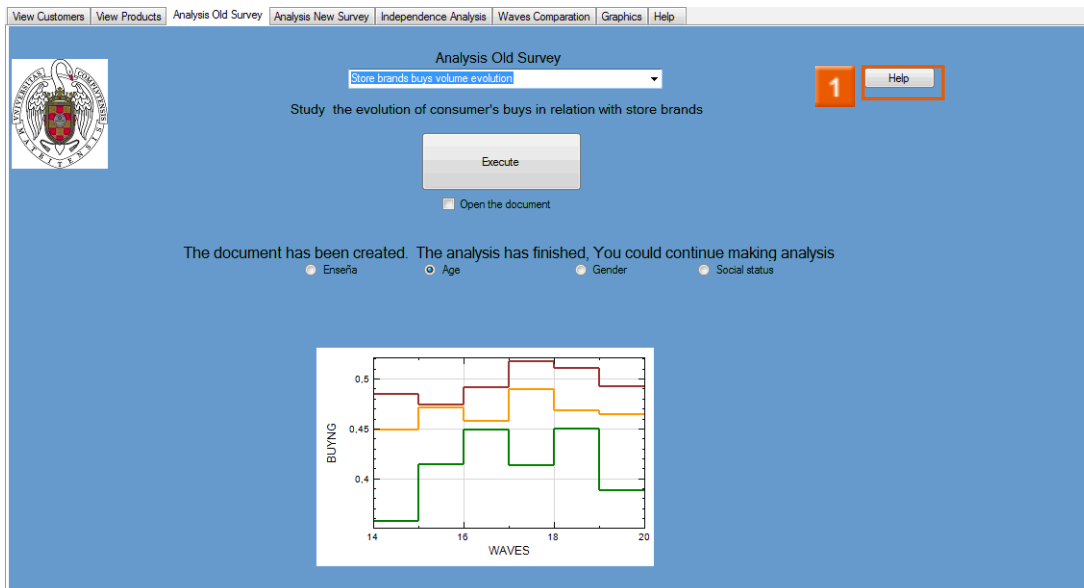
**1** Haga clic en el gráfico. Se mostrará la pestaña *Graphics* con el gráfico ampliado.



**1** Haga clic en **el icono Back** para volver a la pestaña *Analysis Old Survey*.

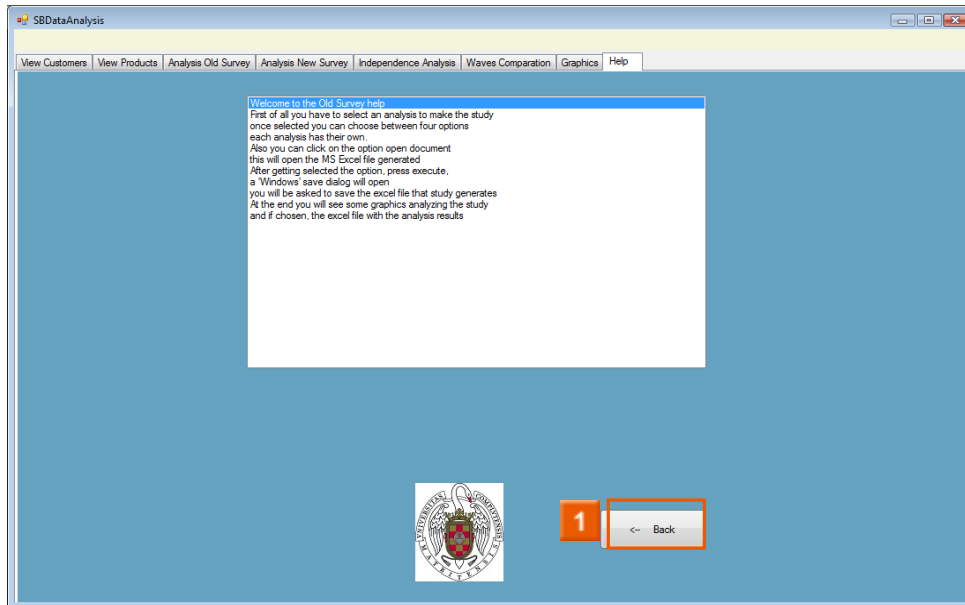
### 3.1.3 Ver ayuda de la pestaña

A continuación se muestra cómo acceder a la ayuda de la ventana donde se muestra el uso de la pestaña.



**1** Haga clic en **el icono Help**. Se mostrará la pestaña *Help* con las instrucciones de uso de la pestaña.



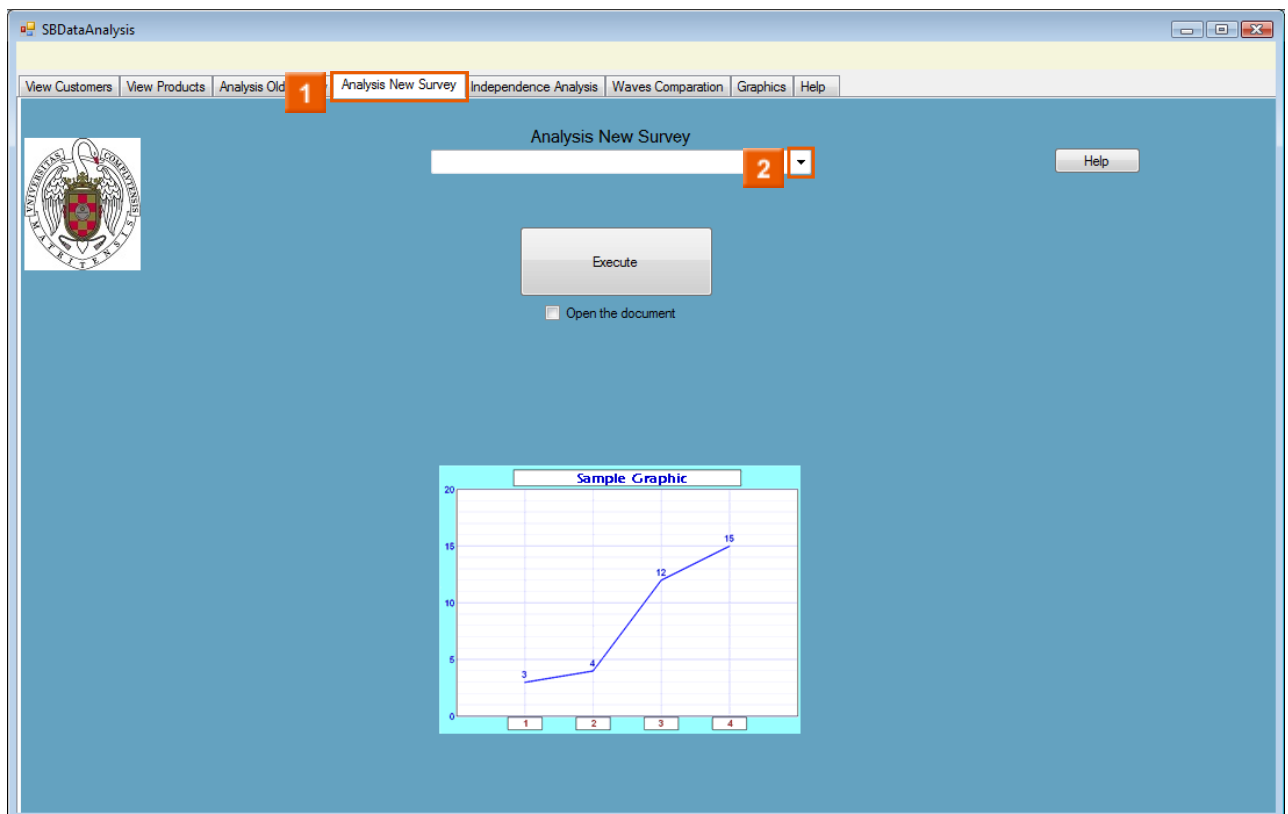


**1** Haga clic en el **icono Back** para volver a las pestaña *Analysis Old Survey*.

## 3.2 Análisis de encuesta de consumidores nueva

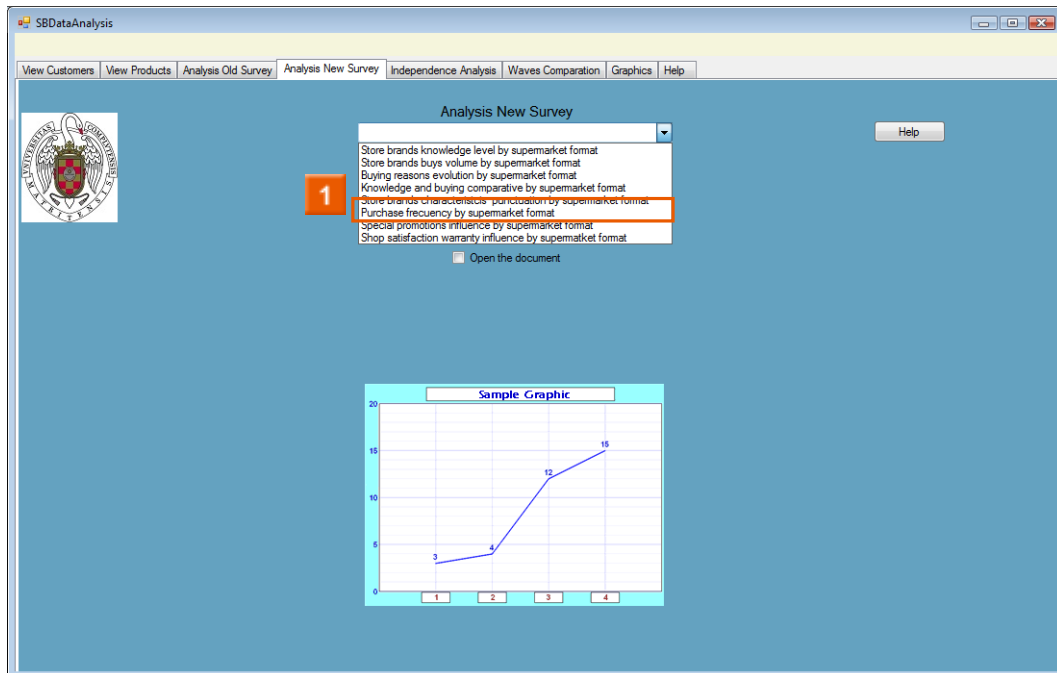
### 3.2.1 Hacer un análisis de las encuestas nuevas a los consumidores

A continuación se explica la realización de un análisis de las encuestas nuevas a consumidores. Se creará un informe en formato Excel y se mostrará unos gráficos con los resultados obtenidos del estudio. En este caso se quiere realizar un análisis de la puntuación a las características de las MDD por formato de enseñanza y estatus social. Además se quiere que al acabar el análisis, se abra el informe con los resultados del análisis.

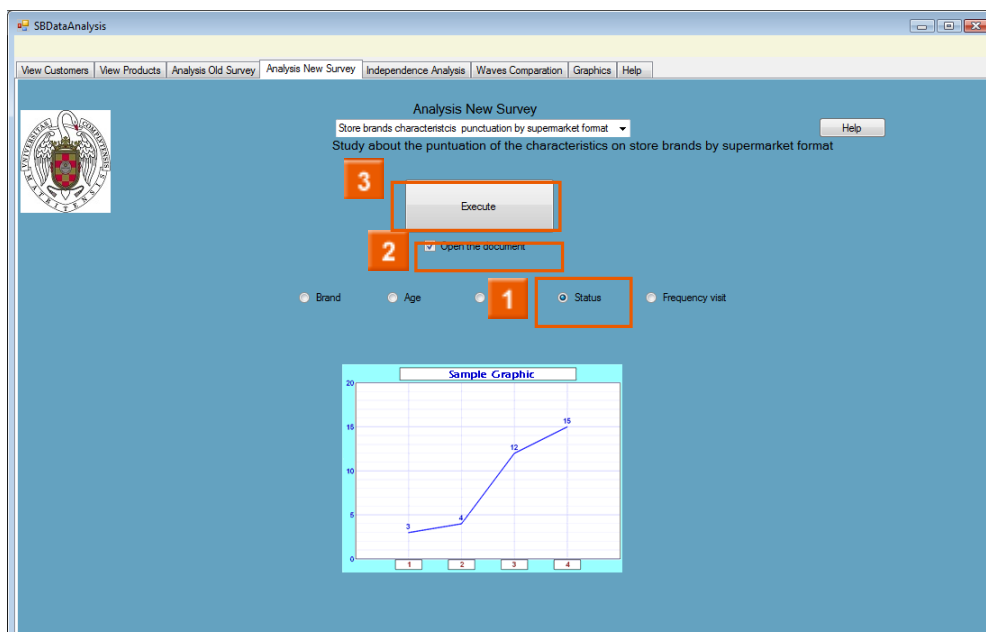


**1** Haga clic en la pestaña **Analysis New Survey**.

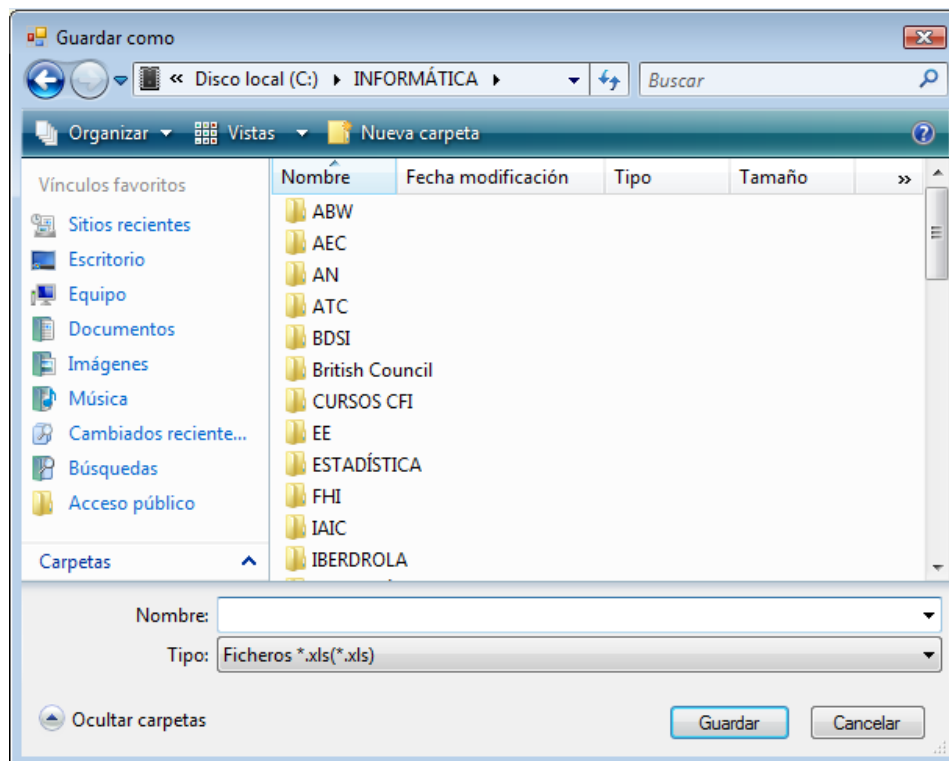
**2** Haga clic en ▼ para la lista de análisis disponibles.



- 1** Haga clic en análisis **Store brands characteristics punctuation by supermarket format**.



- 1** Haga clic en el icono **Status**.
- 2** Haga clic en la casilla **Open the document**.
- 3** Haga clic en el icono **Execute**.



Guarde el informe en la carpeta deseada. Al acabar el análisis, como se ha seleccionado la opción *Open the document* se abrirá el informe y se mostrará un gráfico con los resultados del análisis.



Microsoft Excel - puntuacionCaracteristicas

Archivo Edición Ver Insertar Formato Herramientas Datos Ventana ?

Esciba una pregunta

Arial 10

N A S

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

Sheet1

Lista

Syntax

Resources

EuroValue

desactivar

MEANS TABLES-P17 BY FORMATO  
BY P27 : CELLS COUNT MEAN  
HARMONIC GEOMETRIC VAR  
STDEV MIN MAX KURT SKEW.

0.00 00.015

0.00 00.058

Case Processing Summary

Cases

Included

Excluded

Total

N

Percent

N

Percent

N

Percent

PUNTUACION PRECIO \*  
FORMATO \* STATUS

799

99.6%

3

4%

799

100.0%

Report

PUNTUACION PRECIO

FORMATO

STATUS

N

Mean

Harmonic Mean

Geometric Mean

Variance

Std. Deviation

Minimum

Maximum

HIPER

ACOMODADA

5

.320

.286

.303

.012

.1095

.2

.4

MEDIA ALTA

35

.309

.261

.282

.020

.1491

.2

.4

MEDIA MEDIA

74

.278

.249

.262

.010

.0983

.2

.4

MEDIA BAJA

33

.261

.233

.245

.011

.1059

.2

.4

MODESTA

6

.233

.218

.224

.007

.0816

.2

.4

Total

153

.281

.247

.262

.012

.1111

.2

.4

SUPER

ACOMODADA

13

.338

.284

.309

.023

.1502

.2

.4

MEDIA ALTA

72

.300

.259

.277

.016

.1256

.2

.4

MEDIA MEDIA

142

.286

.250

.266

.014

.1177

.2

.4

MEDIA BAJA

80

.288

.242

.260

.024

.1554

.2

.4

MODESTA

13

.246

.226

.235

.008

.0877

.2

.4

Total

320

.290

.250

.267

.017

.1302

.2

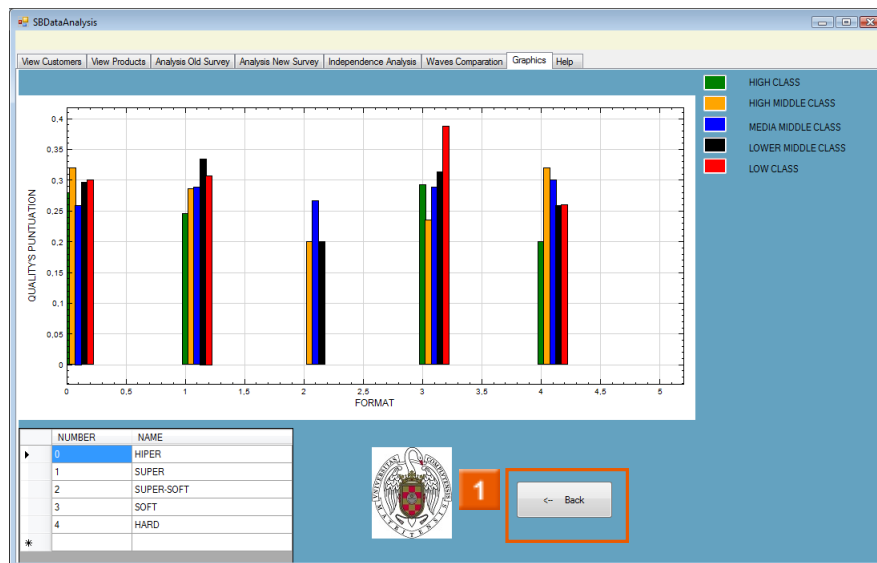
.4

### 3.2.2 Ampliar gráfico de resultado

A continuación se muestra cómo aumentar un gráfico obtenido tras realizar un análisis. Al aumentar el gráfico se verá más claro, además de poder observar su leyenda y el significado de los valores de las abscisas.



**1** Haga clic en el medio del gráfico QUALITY, abajo a la izquierda. Se mostrará la pestaña *Graphics* con el gráfico ampliado.



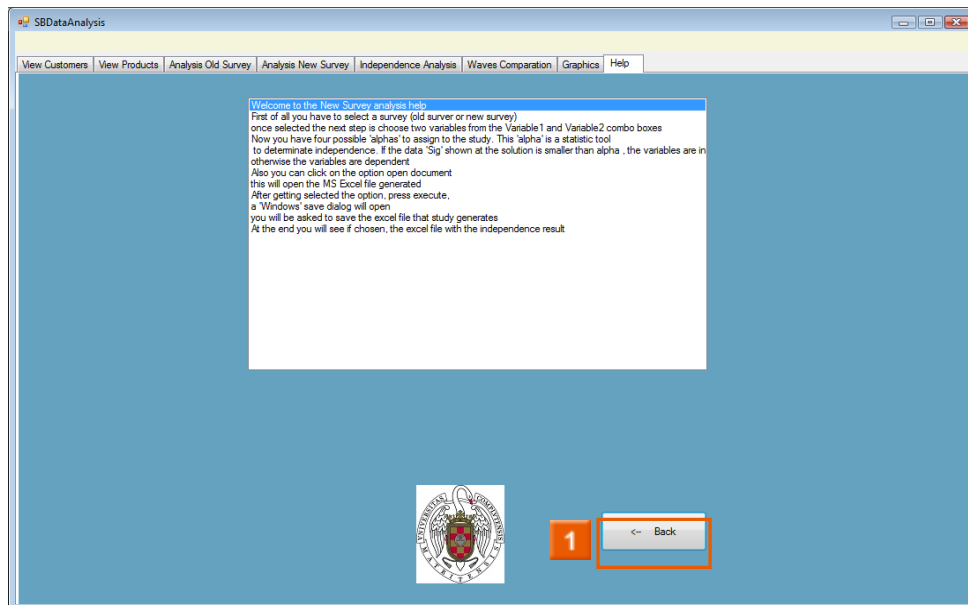
**1** Haga clic en el icono **Back** para volver a la pestaña *Analysis New Survey*.

### 3.2.3 Ver ayuda de la pestaña

A continuación se muestra cómo se puede ver la ayuda de la ventana en la que se explica cómo usar la pestaña.



**1** Haga clic en el icono **Help**. Se mostrará la pestaña *Help* con las instrucciones manual de uso de la pestaña.



**1** Haga clic en el **icono Back** para volver a la pestaña *Analysis New Survey*.

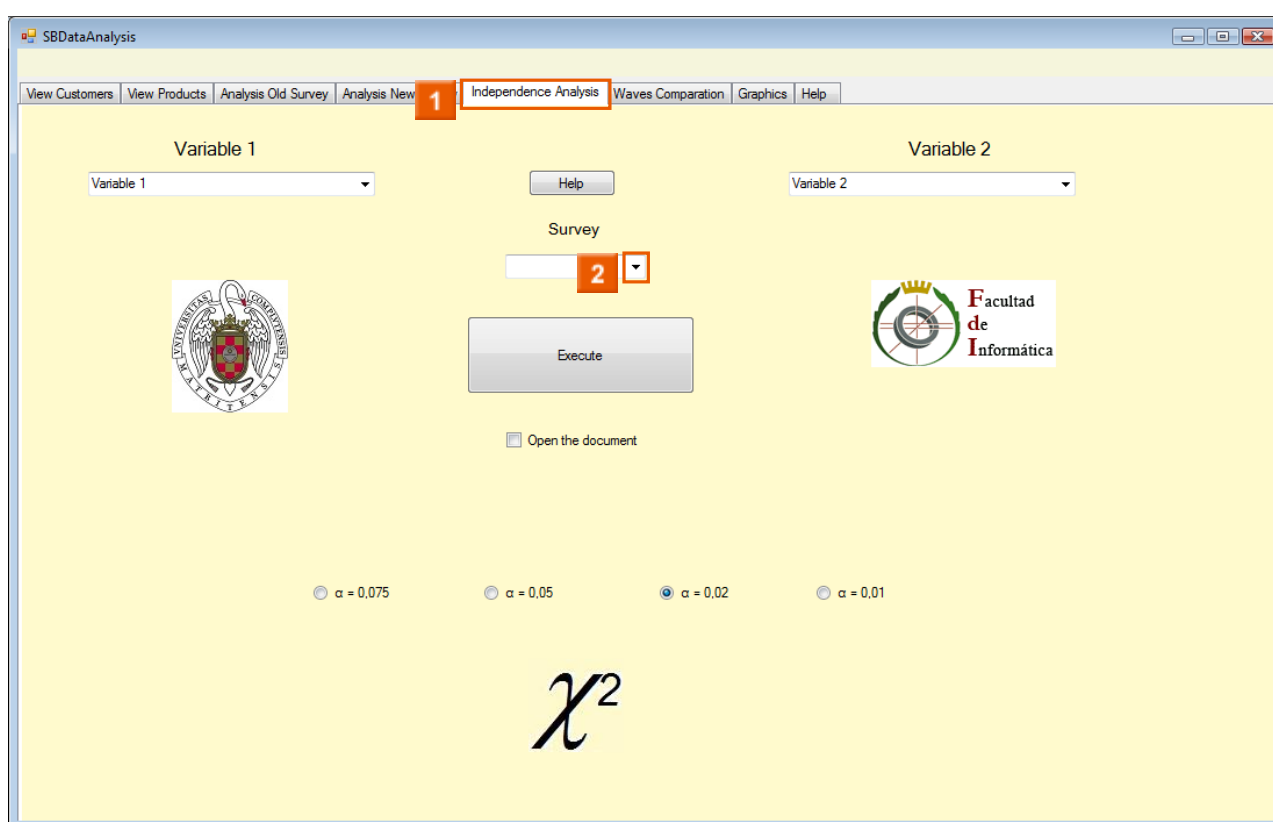





## 4 Análisis de independencia

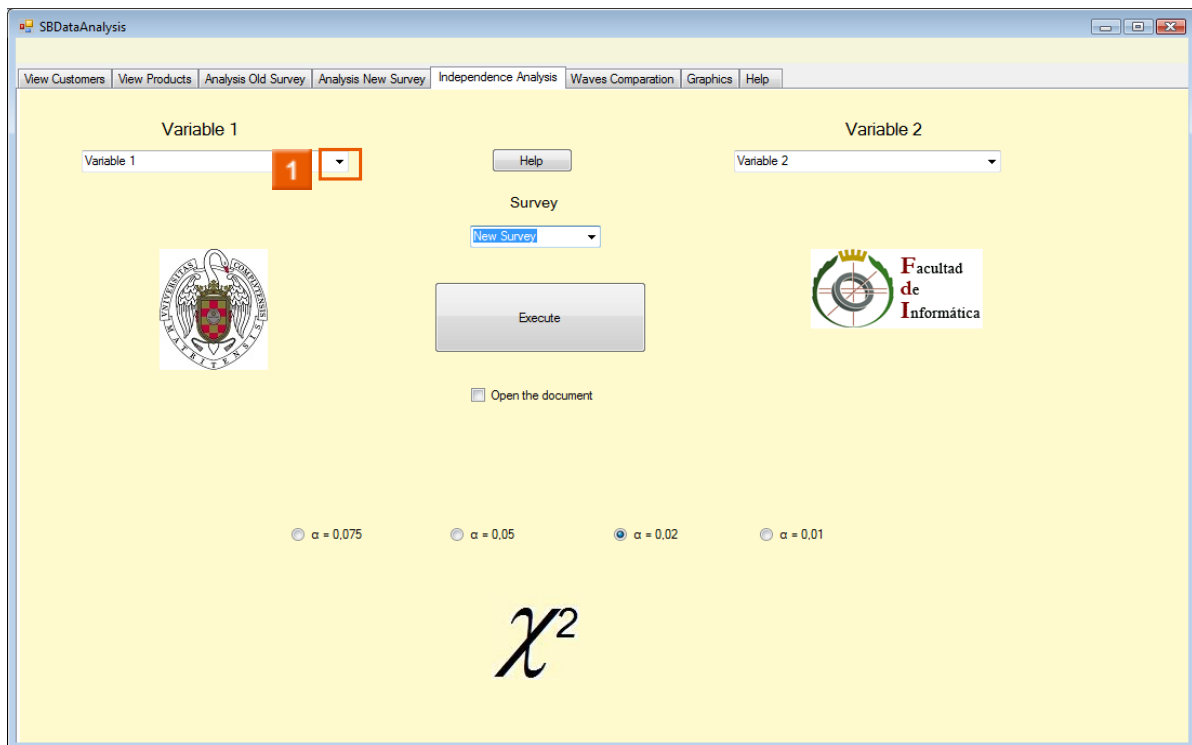
### 4.1 Hacer un análisis de independencia

A continuación se explica cómo realizar un análisis de independencia entre dos variables de las encuestas de consumidores. En este caso se realizará el análisis de independencia a las variables de la encuesta nueva Conocimiento y Puntuación del envase, se usará un alpha de 0,05. Además se quiere que se abra el informe con los resultados, al acabar el análisis.

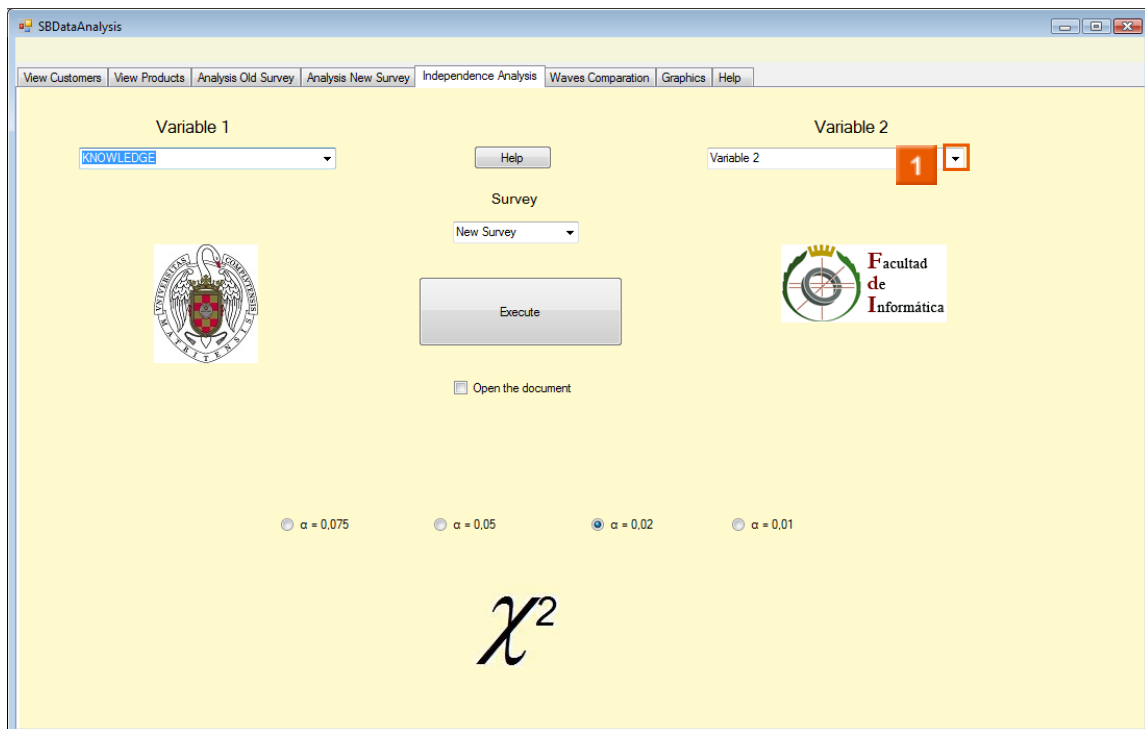


**1** Haga clic en la **pestaña Independence Analysis**.

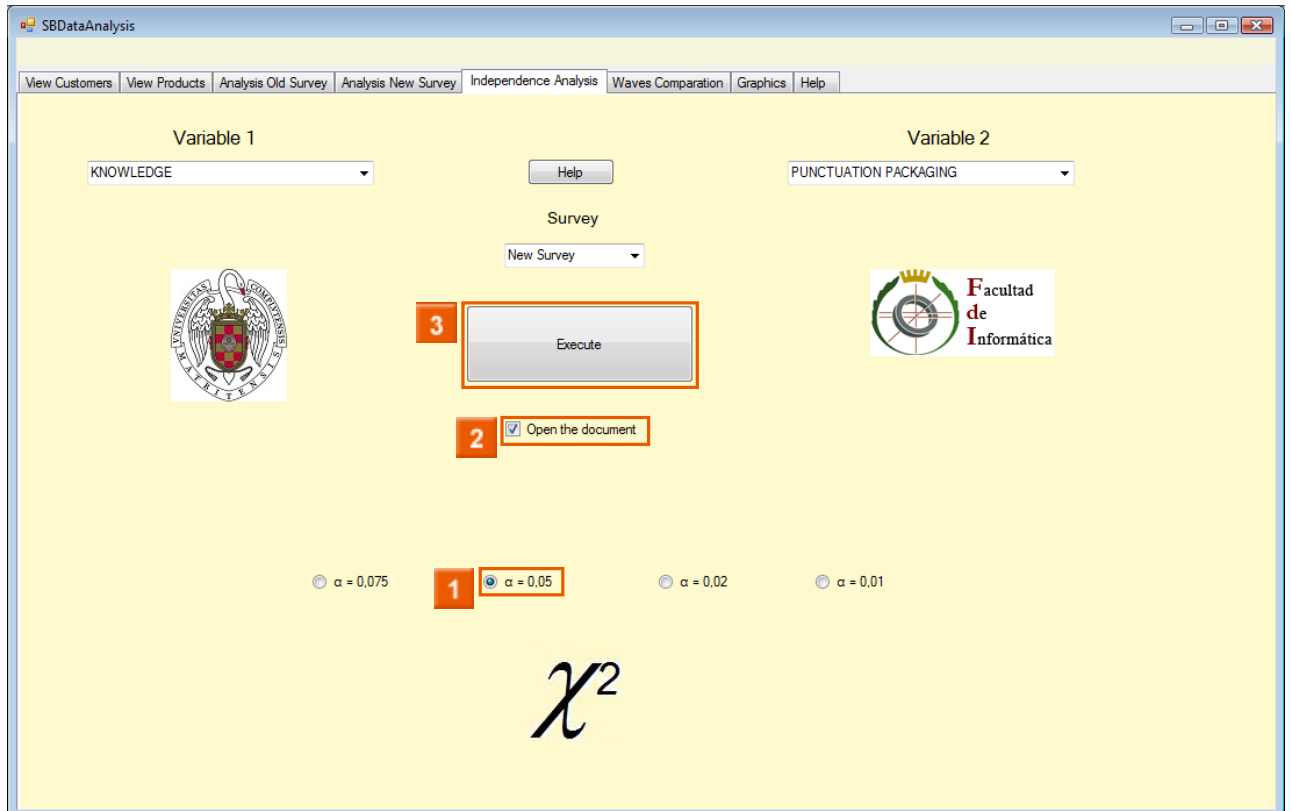
**2** Haga clic en  de la variable Survey y seleccione la opción *New Survey*.



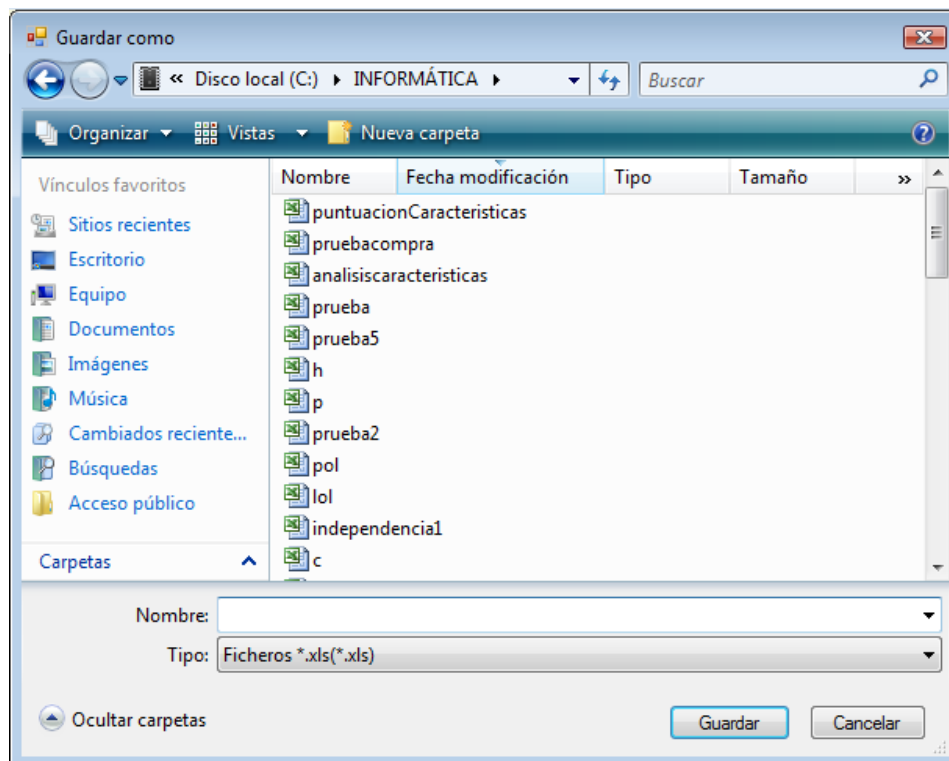
1 Haga clic en ▼ de la Variable 1 y seleccione la variable *KNOWLEDGE*.



1 Haga clic en ▼ de la Variable 2 y seleccione la variable *PUNCTUATION PACKAGING*.



- 1** Haga clic en el icono  $\alpha = 0,05$ .
- 2** Haga clic en la casilla **Open the document**.
- 3** Haga clic en el icono **Execute**.



Guarde el informe en la carpeta deseada. Al acabar el análisis, como se ha seleccionado la opción *Open the document* se abrirá el informe. En el recuadro señalado a continuación, se muestra la tabla del análisis de independencia.

Microsoft Excel - prueba independencia

Archivo Edición Ver Insertar Formato Herramientas Datos Ventana ?

Esciba una pregunta

C40 .033\*.a.b EuroValue desactivar Arial 9

|              |      | PUNTUACION ENVASE |       |       |       |
|--------------|------|-------------------|-------|-------|-------|
|              |      | .2                | .4    | .6    | .8    |
|              |      | Count             | Count | Count | Count |
| CONOCIMIENTO | .00  | 5                 | 6     | 8     | 1     |
|              | .08  | 0                 | 1     | 8     | 0     |
|              | .17  | 2                 | 5     | 10    | 1     |
|              | .25  | 4                 | 16    | 12    | 5     |
|              | .33  | 1                 | 11    | 27    | 2     |
|              | .42  | 11                | 20    | 40    | 3     |
|              | .50  | 3                 | 29    | 34    | 7     |
|              | .58  | 7                 | 18    | 29    | 4     |
|              | .67  | 2                 | 15    | 43    | 3     |
|              | .75  | 6                 | 14    | 23    | 2     |
|              | .83  | 4                 | 14    | 40    | 2     |
|              | .92  | 3                 | 12    | 34    | 1     |
|              | 1.00 | 19                | 64    | 155   | 11    |

Pearson Chi-Square Tests

|              |            | PUNTUACION ENVASE |
|--------------|------------|-------------------|
| CONOCIMIENTO | Chi-square | 53.088            |
|              | df         | 36                |
|              | Sig.       | .033 <sup>a</sup> |

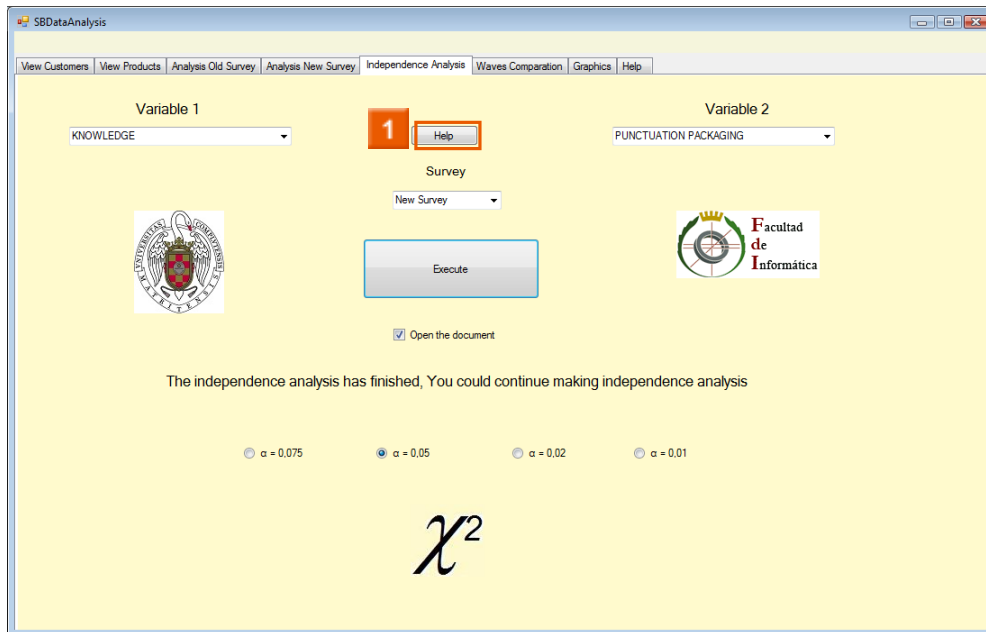
Results are based on nonempty rows and columns in each innermost subtable.

<sup>a</sup>. The Chi-square statistic is significant at the 0.05 level.

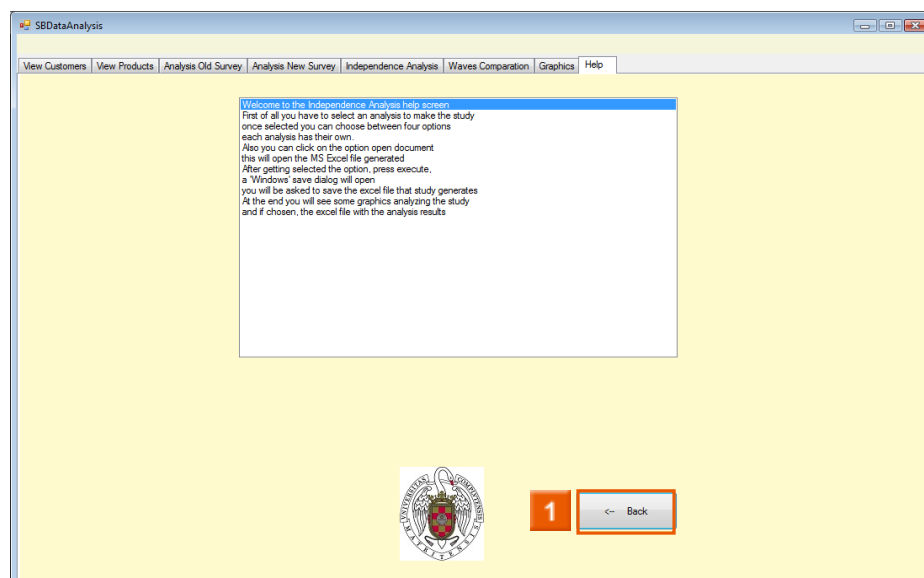
Sheet1

## 4.2 Ver ayuda de la pestaña

A continuación se describe cómo ver la ayuda de la dónde se muestra el uso de la pestaña.



- 1 Haga clic en **el icono Help**. Se mostrará la pestaña *Help* con las instrucciones manual de uso de la pestaña.



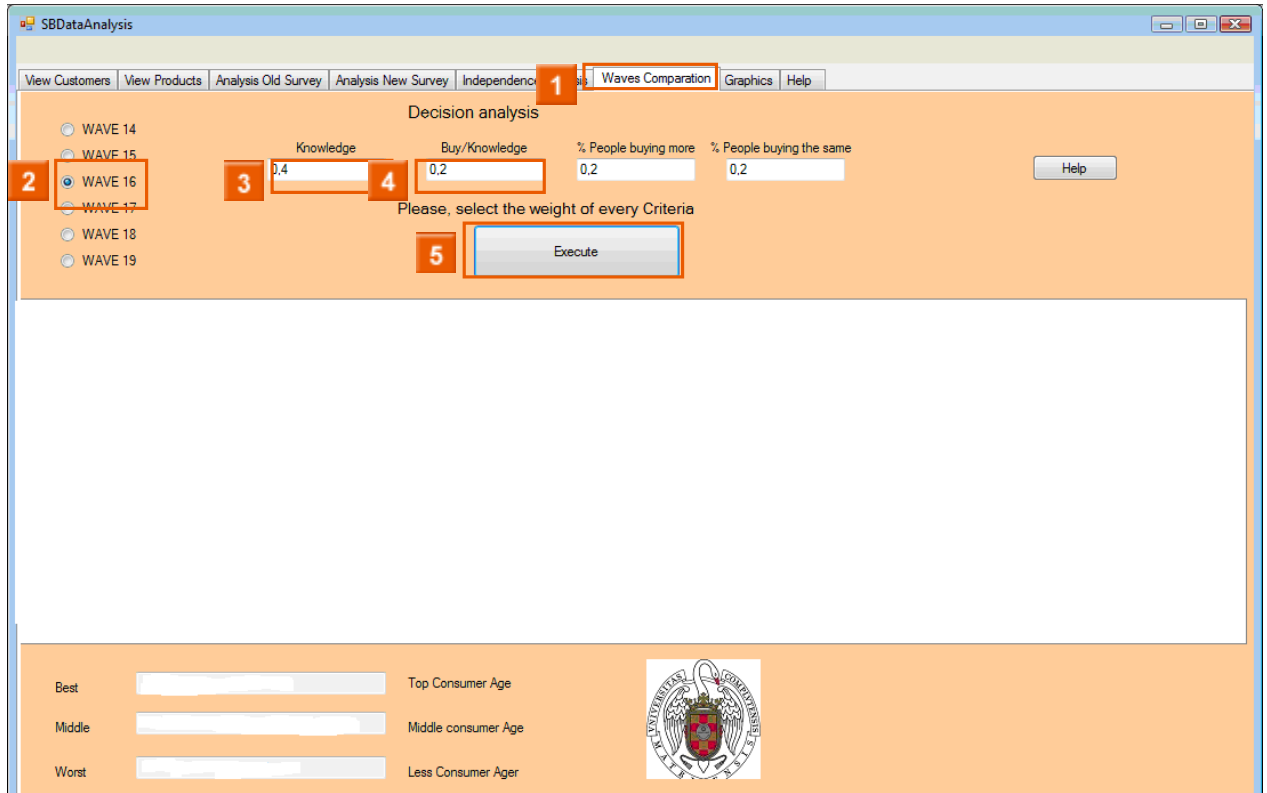
- 1 Haga clic en **el icono Back** para volver a la pestaña *Independence Analysis*.



## 5 Comparación de consumo

### 5.1 Hacer un análisis de comparación de consumo

A continuación se muestra cómo realizar un análisis de consumo utilizando métodos de decisión multicriterio. En este caso se va a realizar con respecto a la Ola 16. Como ponderación de los criterios se usará 0,4 0,2, 0,2 y 0,2 respectivamente.



The screenshot shows the 'SBDDataAnalysis' window with the 'Waves Comparison' tab selected. On the left, a list of waves (WAVE 14 to WAVE 19) is shown, with WAVE 16 selected. The 'Decision analysis' section contains four input fields for criteria weights: 'Knowledge' (0,4), 'Buy/Knowledge' (0,2), '% People buying more' (0,2), and '% People buying the same' (0,2). Below these is a button labeled 'Execute'. At the bottom, there are three rows for consumer age groups: 'Best', 'Middle', and 'Worst', each with a corresponding input field and a label ('Top Consumer Age', 'Middle consumer Age', 'Less Consumer Age'). A university logo is visible in the bottom right corner.

**1** Haga clic en la **pestaña Waves Comparison**.

**2** Haga clic en el **icono WAVE 16**.

**3** Haga clic en la ponderación de Knowledge y introduzca 0,4

**4** Haga clic en la ponderación de Buy/Knowledge y introduzca 0,2

**5** Haga clic en el **icono Execute**. Se mostrará el resultado del análisis y las matrices intermedias calculadas.

SBDataAnalysis

View Customers View Products Analysis Old Survey Analysis New Survey Independence Analysis Waves Comparison Graphics Help

Decision analysis

WAVE 14  
WAVE 15  
☒ WAVE 16  
WAVE 17  
WAVE 18  
WAVE 19

Knowledge Buy/Knowledge % People buying more % People buying the same

0.4 0.2 0.2 0.2

Help

Please, select the weight of every Criteria

Execute

| Knowledge   | % Buy / knowledge  | Buy More   | Buy the same   | range of age                             | kind of matrix              |
|---|--|--|--|--|-----------------------------|
| (0.7 0.78 0.85)<br>(0.67 0.75 0.82)<br>(0.69 0.77 0.84) | (0.52 0.58 0.64)<br>(0.55 0.62 0.68)<br>(0.58 0.64 0.71) | (0.3 0.33 0.36)<br>(0.29 0.32 0.36)<br>(0.19 0.22 0.24)  | (0.39 0.44 0.48)<br>(0.47 0.52 0.57)<br>(0.53 0.59 0.65) | under 25<br>between 25 and 50<br>over 50 | performance matrix          |
| (0.28 0.31 0.34)<br>(0.27 0.3 0.33)<br>(0.28 0.31 0.34) | (0.1 0.12 0.13)<br>(0.11 0.12 0.14)<br>(0.12 0.13 0.14)  | (0.06 0.07 0.07)<br>(0.06 0.06 0.07)<br>(0.04 0.04 0.05) | (0.08 0.09 0.1)<br>(0.09 0.1 0.11)<br>(0.11 0.12 0.13)   | under 25<br>between 25 and 50<br>over 50 | weighted performance matrix |
| 0.69 / 0.31<br>0.7 / 0.3<br>0.69 / 0.31                 | 0.88 / 0.12<br>0.88 / 0.12<br>0.87 / 0.13                | 0.93 / 0.07<br>0.94 / 0.06<br>0.96 / 0.04                | 0.91 / 0.09<br>0.9 / 0.1<br>0.88 / 0.12                  | under 25<br>between 25 and 50<br>over 50 | d* / d- matrix              |

Best Pi= 0.149 -->Over50 Top Consumer Age

Middle Pi= 0.148 -->Between 25 and 50 Middle consumer Age

Worst Pi= 0.145 -->Under 25 Less Consumer Age

## 5.2 Ver ayuda de la pestaña

A continuación se describe cómo ver la ayuda de la ventana en la que se muestra el manual de uso de la pestaña.

SBDataAnalysis

View Customers View Products Analysis Old Survey Analysis New Survey Independence Analysis Waves Comparison Graphics Help

Decision analysis

WAVE 14  
WAVE 15  
☒ WAVE 16  
WAVE 17  
WAVE 18  
WAVE 19

Knowledge Buy/Knowledge % People buying more % People buying the same

0.4 0.2 0.2 0.2

1 Help

Please, select the weight of every Criteria

Execute

| Knowledge   | % Buy / knowledge  | Buy More   | Buy the same   | range of age                             | kind of matrix              |
|---|--|--|--|--|-----------------------------|
| (0.7 0.78 0.85)<br>(0.67 0.75 0.82)<br>(0.69 0.77 0.84) | (0.52 0.58 0.64)<br>(0.55 0.62 0.68)<br>(0.58 0.64 0.71) | (0.3 0.33 0.36)<br>(0.29 0.32 0.36)<br>(0.19 0.22 0.24)  | (0.39 0.44 0.48)<br>(0.47 0.52 0.57)<br>(0.53 0.59 0.65) | under 25<br>between 25 and 50<br>over 50 | performance matrix          |
| (0.28 0.31 0.34)<br>(0.27 0.3 0.33)<br>(0.28 0.31 0.34) | (0.1 0.12 0.13)<br>(0.11 0.12 0.14)<br>(0.12 0.13 0.14)  | (0.06 0.07 0.07)<br>(0.06 0.06 0.07)<br>(0.04 0.04 0.05) | (0.08 0.09 0.1)<br>(0.09 0.1 0.11)<br>(0.11 0.12 0.13)   | under 25<br>between 25 and 50<br>over 50 | weighted performance matrix |
| 0.69 / 0.31<br>0.7 / 0.3<br>0.69 / 0.31                 | 0.88 / 0.12<br>0.88 / 0.12<br>0.87 / 0.13                | 0.93 / 0.07<br>0.94 / 0.06<br>0.96 / 0.04                | 0.91 / 0.09<br>0.9 / 0.1<br>0.88 / 0.12                  | under 25<br>between 25 and 50<br>over 50 | d* / d- matrix              |

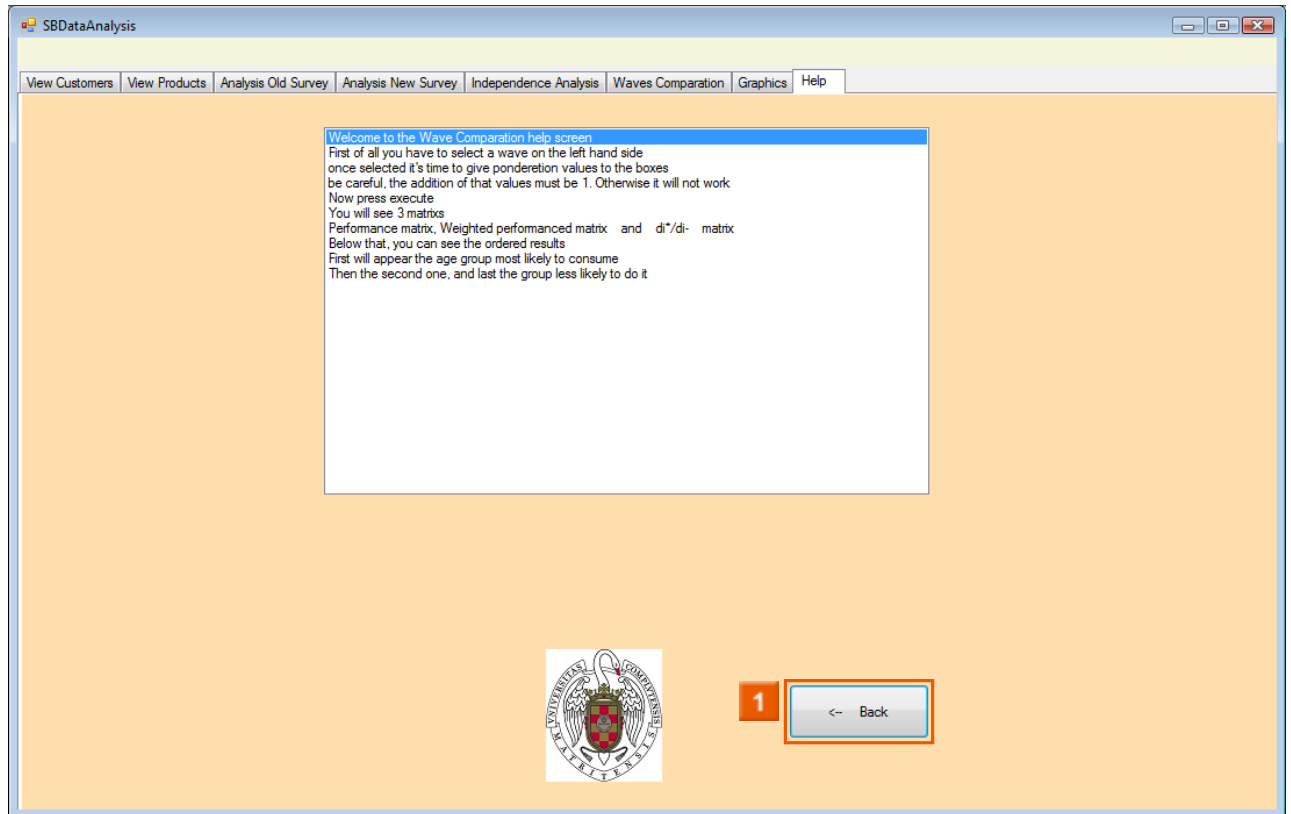
Best Pi= 0.149 -->Over50 Top Consumer Age

Middle Pi= 0.148 -->Between 25 and 50 Middle consumer Age

Worst Pi= 0.145 -->Under 25 Less Consumer Age

**1** Haga clic en **Help**. Se mostrará la pestaña Help con las instrucciones de uso de la pestaña





**1** Haga clic en **el icono Back** para volver a la pestaña *Waves Comparison*.





UNIVERSIDAD COMPLUTENSE DE MADRID



FACULTAD DE INFORMÁTICA

SISTEMAS INFORMÁTICOS

CURSO 2009/2010

ANEXO 2A

# GUÍA DE INVESTIGACIÓN SOBRE MARCAS DE DISTRIBUIDOR

**Jorge Caballero Prados**

**Pablo Rodrigo Jimeno Hidalgo**

**Raúl Ortega Maestro**

*Dirigido por:*

***Dra. María Victoria López López***

*Departamento de Arquitectura de Computadores y Automática*

*Facultad de Informática - Universidad Complutense de Madrid*





INVESTIGACIÓN SOBRE:  
**MARCAS DE DISTRIBUIDOR**  
CURSO 2007 – 2008  
(EDICIÓN VIGÉSIMA) ©PPRGP

Por favor, rellene en primer lugar los siguientes datos:

|   |   |
|---|---|
| <b>- ENSEÑA (*):</b><br><input type="checkbox"/> ALCAMPO<br><input type="checkbox"/> CARREFOUR<br><input type="checkbox"/> EROSKI<br><input type="checkbox"/> EL CORTE INGLES<br><input type="checkbox"/> HIPERCOR<br><input type="checkbox"/> MERCADONA<br><input type="checkbox"/> DIA<br><input type="checkbox"/> MAXI - DIA<br><input type="checkbox"/> SUPERSOL<br><input type="checkbox"/> LIDL<br><input type="checkbox"/> AHORRAMAS | <b>- Nº ESTABLECIMIENTO (*):</b><br>1 0 _ _<br>2 0 _ _<br>3 0 _ _<br>4 0 _ _<br>5 0 _ _<br>6 0 _ _<br>9 0 _ _<br>9 8 _ _<br>1 0 0 _ _<br>1 4 0 _ _<br>1 5 0 _ _ |
| (*)Tachar el que Proceda  | (*) COMPLETAR   |
| <b>DOMICILIO:</b>   |   |
| <b>- Nº EQUIPO:</b>   |   |

Trabajo de Investigación iniciado en 1987 por el Dr. D. José Antonio Puelles, Catedrático de Comercialización e Investigación de Mercados de la Facultad de CC. Económicas y Empresariales de la Universidad Complutense de Madrid (UCM), con la participación de los profesores Ladislao González Ruiz, María Puelles Gallo, Pedro Galván Guijo, Enrique García-Fraile Roca, Miguel Llorens Marín e Ignacio Cuadrado, Carmen Abril Barrié Barrié y Miguel Houghthon, Profesores de la misma Universidad, así como de la Profesora Susana Romero Sanz y el profesor Enrique Pérez del Campo, de la Universidad Rey Juan Carlos y el Profesor Rafael Pascual Camino de la Universidad Pontificia de Salamanca.



**MUY IMPORTANTE:**

- Sólo se acudirá al establecimiento después de haber leído y haber comprendido esta GUIA DE LA INVESTIGACIÓN.**
- Para rellenar los cuestionarios electrónicos será imprescindible utilizar una versión de MICROSOFT EXCEL XP O MICROSOFT EXCEL 2003, ya que en versiones anteriores no funciona correctamente.**
- Será necesario activar la seguridad de macros, ya que los archivos incorporan programas que facilitan el registro de los datos. Estas macros no contienen virus y son inofensivas para los ordenadores. Ver instrucciones en los cuestionarios electrónicos.**
- Para guardar los datos de cada centro investigado, se abrirá una carpeta por cada uno de los establecimientos, consignado su número y la enseña, p.ej.: 2001 CARREFOUR.**

El presente documento constituye el soporte informativo, en la parte correspondiente a la GUIA DE LA INVESTIGACIÓN, como el soporte para la TOMA DE DATOS de la investigación.

El contenido de este documento es el siguiente:

1. GUIA DE LA INVESTIGACIÓN:

- INSTRUCCIONES ACERCA DE LOS ESTABLECIMIENTOS.
- CALENDARIO Y ALCANCE DEL TRABAJO

2. CUESTIONARIO DE COMPRADORES:

- INSTRUCCIONES DEL CUESTIONARIO DE COMPRADORES.

3. CUESTIONARIO DE PRODUCTOS:

- INSTRUCCIONES DEL CUESTIONARIO DE PRODUCTOS.
- **1 FORMULARIO PARA RELLENAR.**



# 1. GUIA DE LA INVESTIGACIÓN

## **INSTRUCCIONES SOBRE LOS ESTABLECIMIENTOS:**

El objetivo central y principal de la investigación es la recogida, depuración y estudio de los datos correspondientes a productos comercializados con *Marcas del Distribuidor*, hecho este que debe prevalecer en el enfoque del trabajo. El hecho de recoger datos referentes a productos comercializados con Marcas de Fabricantes es a efectos comparativos y su concurrencia es por tanto necesaria.

Por tanto, se deberá poner mayor énfasis, esfuerzo y cuidado a la hora de tomar datos de productos MD, a la hora de redactar los informes, enfocando la redacción a estos informes, y en general a todas las actividades que deberán realizar en este trabajo y que se detallan a continuación en esta Guía de la Investigación.

A los efectos de esta investigación, marca blanca es igual a marca de distribuidor y se designará por **MD**, mientras que el resto, es decir, marcas no blancas, marcas de fabricante o marcas nacionales, se designarán en esta guía como **MF**.

La atribución de los establecimientos se hace a cada equipo, siendo este el que a su vez realice la distribución a cada uno de los componentes. Las condiciones son que cada miembro cumplimente al menos un cuestionario completo y el equipo alcance **1,5 puntos por componente** con la siguiente distribución:

| <b>ESTABLECIMIENTO</b> | <b>VALOR</b> |
|------------------------|--------------|
| ALCAMPO                | 4 puntos     |
| EL CORTE INGLÉS        | 4 puntos     |
| HIPERCOR               | 4 puntos     |
| EROSKI                 | 3 puntos     |
| LIDL                   | 3 puntos     |
| AHORRAMAS              | 3 puntos     |
| MERCADONA              | 2 puntos     |
| SUPERSOL               | 2 puntos     |
| MAXI DIA               | 2 puntos     |
| DIA                    | 1 punto      |

Conviene señalar que la calidad del trabajo dependerá de cada investigador, por lo que el mal trabajo de un investigador afectará a la calidad del conjunto de la investigación.

En cada cuestionario en papel, se consignará siempre el nombre de la enseña (Carrefour, Día, Hipercor, etc.), el nº del establecimiento (4 cifras), la dirección del establecimiento objeto de estudio, y el nº de GRUPO asignado, al objeto de identificar correctamente cada uno.



### **CALENDARIO Y ALCANCE DEL TRABAJO.**

1. La captura de datos en los establecimientos se realizará utilizando los dos cuestionarios de papel o manuales (de productos y de comprador), sobre los que se escribirá con lápiz o bolígrafo.
2. La captura de datos deberá realizarse **EXCLUSIVAMENTE ENTRE: el lunes 26 de noviembre y el domingo 9 de diciembre** del presente año 2007 (ambos inclusive).
3. Posteriormente se introducirán en los soportes informáticos todos los datos anteriormente capturados en el cuestionario de papel, siguiendo las instrucciones de los Cuestionarios Excel.
4. Finalmente, establecimiento por establecimiento, se preparará un informe en el que se resalten los aspectos más interesantes referidos a aspectos tales como los siguientes (siendo la relación meramente enunciativa y no exhaustiva):
  - Opinión y disposición de los productos y zonas frías y calientes en el establecimiento.
  - Aspectos relevantes de las acciones de merchandising en el establecimiento.
  - Disposición del establecimiento y organización del tráfico de clientes.
  - Aspectos positivos y negativos para cada establecimiento engeneral, y en relación con las MD en particular.
  - Dificultades surgidas en relación con la realización de las encuestas.
  - Anécdotas vividas en la realización del trabajo de campo.
  - Aspectos positivos y negativos de la investigación y sugerencias para mejorar la investigación en las próximas ediciones.
  - Se podría incorporar cualquier información (tomada directamente de Internet incluso) sobre algunas de las enseñanzas analizadas o sobre el fenómeno de MD.

Cada equipo entregará la siguiente documentación **de cada establecimiento:**

1. Cuestionario en papel, con los datos tomados a mano.
2. Archivos Excel con los datos transcritos y los informes de cada establecimiento, grabados en el CD vacío entregado a cada equipo.

El trabajo terminado se entregará como máximo en la primera semana del mes de enero de 2008, **entre el lunes 07 y el lunes 14 de enero de 2008.**





## **2. CUESTIONARIO DE COMPRADORES.**

Se cumplimentará **entrevistando a un comprador o compradora**. La entrevista podrá realizarse dentro del propio establecimiento (lo ideal) o fuera a la salida de este (comprobando que ha realizado alguna compra, porta bolsa del establecimiento).

El entrevistador deberá limitarse a **recoger la opinión de la persona entrevistada y en ningún caso la propia. Sí no hay opinión se utilizara la casilla "no contesta" ò "no sabe"**.

El mejor sistema para realizar la entrevista es hacerla por parejas. Donde uno de los entrevistadores haga de captador, es decir seleccione las personas a entrevistar, de tal forma que no siempre se seleccione a un mismo tipo de comprador. Y, asegurándonos que, efectivamente, es un/a reciente comprador/a de ese centro:

- Presentándose correctamente
- Informado, previamente, del objeto de la entrevista.
- Requiriendo su colaboración
- Y asegurándonos de que su predisposición es totalmente favorable para someterse al cuestionario.
- Y muy cortésmente se la presenta al compañero que esta realizando la entrevista.

Una vez que la persona comienza a ser entrevistada por la segunda persona, la persona que realice la captación comenzara de nuevo la selección del próximo/a comprador/a a entrevistar.

El otro componente del equipo pasará inmediatamente a cumplimentar el cuestionario, leyendo pausadamente cada pregunta y asegurándonos de que la respuesta dada por la entrevistada corresponde exactamente a lo anotado en el cuestionario.

**Se cumplimentarán al menos 3 encuestas por cada centro estudiado, de forma que se considerará positivamente la aportación de un mayor número de encuestas.**

Para completar el perfil socioeconómico, se utilizará la codificación que aparece en un cuadro al final de cada encuesta.



## **ACLARACIONES AL CUESTIONARIO INFORMÁTICO**

Para activar las macros, abrir Excel, en el menú:  
Herramientas/Macros/Seguridad de Macros/ nivel medio.  
Cerrar Excel y volver a abrir, y cuando pregunte si habilita macros,  
pulsar la opción "Habilitar macros", el programa funcionará entonces  
correctamente.

**Al abrir el archivo aparecerán cuadros que preguntarán: el N° de ESTABLECIMIENTO, el N° de GRUPO, así como FECHA Y HORA. Cuando se introducen estos datos, el programa graba automáticamente el archivo con el código de cada centro que hayan escrito en el cuadro del N° de centro el n° del establecimiento y la letra correspondiente a la encuesta:  
p.ej. 8026A**

\* **Pregunta 1:** Respuesta única, posibles: **SI** o **NO** (anotar siempre una de las 2 posibilidades con un X, S o N)

\* **Pregunta 2:** Rellenar en el cuadro la respuesta a cada pregunta solamente en caso afirmativo **con [X o S]**. En caso negativo, no anotar nada.

\* **Pregunta 3:** Seleccionar, una o más de las siguientes respuestas, anotando en la/s casilla/s correspondiente/s con **[X o S]**:

\* **Preguntas 4 y 5. Respuesta única.** Tienen las siguientes opciones:

|                |                  |                  |                    |                        |
|----------------|------------------|------------------|--------------------|------------------------|
| <b>1 - Más</b> | <b>2 - Menos</b> | <b>3 - Igual</b> | <b>4 - No sabe</b> | <b>5 - No contesta</b> |
|----------------|------------------|------------------|--------------------|------------------------|

Rellenar la celda correspondiente **con [X o S]**.

\* **Pregunta 6. Respuesta única.** En cada apartado se pondrá:

|                           |                       |                            |                    |                        |
|---------------------------|-----------------------|----------------------------|--------------------|------------------------|
| <b>1 - Muy importante</b> | <b>2 - Importante</b> | <b>3 - Poco importante</b> | <b>4 - No sabe</b> | <b>5 - No contesta</b> |
|---------------------------|-----------------------|----------------------------|--------------------|------------------------|

Rellenar la celda correspondiente por cada fila **con [X o S]**

\* **Preguntas 7, 8 y 9. Respuesta única:**

|              |               |                    |                        |
|--------------|---------------|--------------------|------------------------|
| <b>1 -SI</b> | <b>2 - NO</b> | <b>3 - No sabe</b> | <b>4 - No contesta</b> |
|--------------|---------------|--------------------|------------------------|

Rellenar la celda correspondiente **con [X o S]**

\* **Pregunta 10. Respuesta única por fila:**

|                  |                       |                 |                    |                        |
|------------------|-----------------------|-----------------|--------------------|------------------------|
| <b>1 - Mucha</b> | <b>2 - Importante</b> | <b>3 - Poca</b> | <b>4 - No sabe</b> | <b>5 - No contesta</b> |
|------------------|-----------------------|-----------------|--------------------|------------------------|

Rellenar la celda correspondiente por cada fila **con [X o S]**



- **Pregunta 11.**

DATOS DE IDENTIFICACIÓN CON RESPUESTA ÚNICA:

- × EDAD: Rellenar la celda correspondiente por cada fila **con [X o S]**.
- × SEXO: Rellenar la celda correspondiente **con [X o S]**.
- × LUGAR DE RESIDENCIA: Rellenar la celda **con [X o S]**.

DATOS DE IDENTIFICACIÓN CON DATOS COMO RESPUESTA:

- × Nº PERSONAS QUE VIVEN EN EL HOGAR: Rellenar con el dato.
- × FRECUENCIA DE LA COMPRA: Rellenar con el dato.
- × STATUS SOCIOECONÓMICO: Una vez seleccionado en el cuadro que figura al final de la encuesta, rellenar la celda **con [X o S]**.

Una vez consignados todos los datos en las celdas correspondientes, deberán realizar una sencilla comprobación, pulsando el botón que figura en la parte superior de la hoja: **COMPROBACIÓN FINAL**

Este botón marcará en color rojo aquel grupo de celdas que, siendo obligatorio una respuesta única, se hayan consignado más de una.

En ese caso, deberán de corregir la anomalía y volverán a pulsar el botón **COMPROBACIÓN FINAL**, hasta que no aparezca ninguna celda en color rojo.

Una vez realizada esta verificación, pueden proceder a GRABAR, para después imprimir la encuesta. EL ARCHIVO YA ESTÁ CONFIGURADO PARA IMPRESIÓN, por lo que no necesitarán realizar ningún ajuste, simplemente imprimir.



### **3. CUESTIONARIO DE PRODUCTOS**

El cuestionario de productos debe rellenarse con los datos obtenidos de la observación in-situ de las distintas variables que van a ser objeto de estudio: Producto, Precio, Promoción y Emplazamiento. En total serán 157 artículos, que aparecerán consecutivamente, y que están agrupados por tipos de productos: lácteos, conservas, pastas, etc.

Cada uno de los 157 productos ha sido definido como un producto concreto, en un tipo de envase y con una capacidad delimitada. Así definido el producto, será una REFERENCIA.

**Se investigará un producto siempre y cuando aparezca como MD, en ese caso se investigarán adicionalmente hasta 6 MF para la misma referencia, aquellas a las que la enseña dedique más espacio en el establecimiento.**

**Si aparecen más de una MD, se pondrán todas las que aparezcan (Por ejemplo Grupo Carrefour: Carrefour, Champion, Dia).**

**En todos los productos se facilitan una o varias de las marcas líderes en MF al objeto de servir de referencia, teniendo en cuenta que puede que estén o no, todas o en parte, en los establecimientos.**

Las tres primeras filas se reservarán para MD y las restantes para las MF, comenzando a consignar estas, por tanto, a partir de la cuarta fila, rellenando en cada caso todas las columnas..

#### **INSTRUCCIONES DEL CUESTIONARIO DE PRODUCTOS**

Para activar las macros, abrir Excel, en el menú:  
Herramientas/Macros/Seguridad de Macros/ nivel medio.  
Cerrar Excel y volver a abrir, y cuando pregunte si habilita macros, pulsar la opción "Habilitar macros", el programa funcionará entonces correctamente.

A continuación indicaremos el contenido de cada una de las columnas, como información de cada referencia. Al ser el mismo formato en papel y en el cuestionario electrónico, los datos capturados en papel se transcribirán íntegros a la tabla de datos de Excel.

|  |
|--|
| <b>Al abrir el archivo MdCuestionarioProductosOla19, se preguntará el Nº de centro (en número). Cuando se introducen estos datos, el programa graba automáticamente el archivo con el código de cada centro que hayan escrito en el cuadro del Nº de centro:</b> |
|--|



|        |    |                             |          |
|--------|----|-----------------------------|----------|
| Centro | NP | Descripción del<br>Producto | MD<br>MF |
|--------|----|-----------------------------|----------|

Por columnas:

- **Columna A [CENTRO]:** Se autorellena al introducir el N° de centro.
- **Columna B [NP]:** Ya relleno. Indica el código de producto.
- **Columna C [DESCRIPCIÓN DEL PRODUCTO]:** Recoge las características del producto que va a ser objeto de estudio, tanto de las genéricas (leche), como las específicas (UHT), (Natural), así como el tipo de envase (BRIK), y el peso o capacidad (1 L). (REFERENCIA)
- **Columna D [MD/MF]:** Indica que en la fila, debe indicarse la información referente a cada tipo de producto, en letra normal para MD y en letra *KURSIVA* MF.

A PARTIR DE ESTAS, HAY QUE INTRODUCIR LOS DATOS:

|       |            |
|-------|------------|
| MARCA | FABRICANTE |
|-------|------------|

- **Columna E [MARCA]:** En esta columna se ha de indicar la MARCA con la que el producto MD o el MF se presenta para la venta. Por ejemplo, en MD de Carrefour la marca puede ser *Carrefour* o "*De nuestra tierra*", o en MD de Mercadona la marca *Hacendado*.
- **Columna F [FABRICANTE].** En esta columna se ha de indicar el nombre de la empresa que fabrica el producto. Hay que consignar el nombre en su forma más conocida: PASCUAL DE ARANDA, SA – PASCUAL. En su defecto, los 7 primeros dígitos del código de barras.

| A COMPARAR   |         |
|--------------|---------|
| N Uds<br>Std | Uds Std |

- **Columnas G y H [N° UDS STANDARD Y UDS STANDARD]:** En estas dos celdas se indica la unidad de medida que tienen que comparar y la magnitud: 300 gramos, deben escribir a continuación el valor real, por ejemplo 280 gramos



|                                 |                |                |                  |
|---------------------------------|----------------|----------------|------------------|
| PESO /<br>CAP /<br>REAL /<br>UD | Nº UDS<br>REAL | PRECIO<br>REAL | PRECIO<br>UNIDAD |
|---------------------------------|----------------|----------------|------------------|

- **Columna I [PESO/CAP/REAL].** En las columna G y H aparecen una descripción de todos y cada uno de los productos objeto de estudio. El Peso o capacidad del producto que pone a la venta el distribuidor puede variar por muchos motivos.  
En esta columna se ha de insertar el peso real, el que indica el producto, **siendo admisible aquellos cuya variación sea superior o inferior en un 25% del indicado** en dicha columna. Por ejemplo, yogur de 115 grs.  
**En esta celda se deberá escribir según las UDS del producto descrito.**
- **Columna J [Nº UDS REAL].** Al igual que en el caso anterior, se indicará el nº de uds en que está presentado el producto. Por ejemplo, la cuajada puede venir presentada en formato de 2 UDS, o en formato de 4 UDS.
- **Columna K [PRECIO REAL].** Es el precio de la unidad o del formato de varias unidades. Por ejemplo, un litro de leche presenta un precio unitario, sin embargo, la cuajada presenta un precio para 2 UDS o bien para 4 UDS, según el formato.
- **Columna L [PRECIO UNIDAD]. CELDA NO MODIFICABLE.** El programa autocalcula en función de las tres variables anteriores el precio correspondiente a una unidad, que será el objeto de estudio.

## LIDER

- **Columnas M [LIDER]. RESPUESTA ÚNICA. Se indicará en esta columna la MARCA LIDER de entre las presentes.**  
A efectos de esta investigación entenderemos como líder a la marca líder natural, o bien a aquella con mayor precio.
- **Columna N,O,P [ENVASE]. RESPUESTA ÚNICA, SOLO PARA MD.** Se indicará en esta celda si el envase de la marca MD objeto de estudio es Mejor, Igual o Inferior que el de la marca Líder de MF. Se consignará la celda correcta, según corresponda.



- **Columnas Q,R,S [ESPACIO]. RESPUESTA ÚNICA.** Se indicará en esta celda si el espacio estimado que se dedica a la marca objeto de estudio es superior a 1/3 del espacio total dedicado al producto, si es 1/3 o bien si es inferior a 1/3, consignando la respuesta correcta en una única celda, con una **X**.
- **Columnas T, U, V, W [COLOCACION/LINEAL].** Según se halle en uno o varios de los posible sitios en los que ubicarse en el lineal, los casos posibles son los siguientes, consignando las correspondientes:

|              |
|--------------|
| Altura ojos  |
| Altura manos |
| Altura techo |
| Altura suelo |

**EN LIDER, ENVASE Y LINEAL, LA RESPUESTA SOLO PUEDE SER ÚNICO, SINO ES ASÍ, TRAS EL CHEQUEO, APARECERÁN EN COLOR ROJO.**

- **Columnas X, Y, Z, AA [COLOCACION/DESTACADOS].** Según se halle en uno o varios de los posibles sitios **fuera del lineal**, los casos posibles son los siguientes, consignando las celdas correspondientes con una X:

|                    |
|--------------------|
| Cabecera de lineal |
| Isla               |
| Chimenea           |
| Otros              |

- **Columnas AB-AG [PROMOCION].** Serán objeto de estudio aquellos estímulos de merchandising y PLV que contribuyan a destacar al producto del resto EN EL PROPIO ESTABLECIMIENTO. Los más genéricos y por tanto objeto de estudio son los siguientes:

|         |   |          |               |       |                      |
|---------|---|----------|---------------|-------|----------------------|
| Folleto | Publicidad centro (tv, video, megafonia, etc) | Carteles | degustaciones | Otras | <b>Sin promocion</b> |
|---------|---|----------|---------------|-------|----------------------|

Se consignarán las celdas correspondientes con una X.



Una vez consignados todos los datos en las celdas correspondientes, deberán realizar una sencilla comprobación, pulsando el botón que figura en la parte superior de la hoja: **COMPROBACIÓN FINAL**, este proceso durará unos segundos o unos minutos en función del equipo informático que lo ejecute.

Este botón marcará en color rojo aquel grupo de celdas que, siendo obligatorio una respuesta única, se hayan consignado más de una.

En ese caso, deberán de corregir la anomalía y volverán a pulsar el botón **COMPROBACIÓN FINAL**, hasta que no aparezca ninguna celda en color rojo.

Una vez realizada esta verificación, pueden proceder a imprimir la encuesta. EL ARCHIVO YA ESTÁ CONFIGURADO PARA IMPRESIÓN, por lo que no necesitarán realizar ningún ajuste, simplemente imprimir.





**UNIVERSIDAD COMPLUTENSE DE MADRID**



**FACULTAD DE INFORMÁTICA**

**SISTEMAS INFORMÁTICOS**

**CURSO 2009/2010**

# **ANEXO 2B**

## **ENCUESTA A CONSUMIDORES**

**Jorge Caballero Prados**

**Pablo Rodrigo Jimeno Hidalgo**

**Raúl Ortega Maestro**

*Dirigido por:*

***Dra. María Victoria López López***

*Departamento de Arquitectura de Computadores y Automática*

*Facultad de Informática - Universidad Complutense de Madrid*





## GRUPO GUIA-MDD OLA 22

CENTRO:

Nº EQUIPO:

DÍA:

HORA:

Hola, buenos días / tardes.

Desde hace 22 años la Universidad Complutense de Madrid e ICADE, realizan un estudio sobre Marcas de Distribuidor (también conocidas como Marcas Blancas), que son productos con algunas de las marcas propias y exclusivas de este establecimiento en el que acaba de comprar Ud., (decir alguna de ellas, si procede, dependiendo del establecimiento).

Creemos que este tema le resultará interesante, y sus respuestas nos pueden ayudar al estudio de este fenómeno. MUCHAS GRACIAS

P1: Del siguiente listado, díganos cuales son las principales razones para elegir este supermercado. Puede elegir varias razones. [marcar las que señale el encuestado]:

[Respuesta múltiple]

|                                 |                          |                                  |                          |                                    |                          |
|---------------------------------|--------------------------|----------------------------------|--------------------------|------------------------------------|--------------------------|
| Precio                          | <input type="checkbox"/> | Muchas marcas de cada producto   | <input type="checkbox"/> | Servicio a domicilio               | <input type="checkbox"/> |
| Cercanía                        | <input type="checkbox"/> | Sus Marcas de distribuidor       | <input type="checkbox"/> | Limpieza y aspecto agradable       | <input type="checkbox"/> |
| Variedad de productos           | <input type="checkbox"/> | Calidad de los productos frescos | <input type="checkbox"/> | Tiene otras tiendas de ocio, ropa, | <input type="checkbox"/> |
| Confianza en el establecimiento | <input type="checkbox"/> | Ofertas/ promociones             | <input type="checkbox"/> | No es sólo de alimentación         | <input type="checkbox"/> |
| Buen trato al cliente           | <input type="checkbox"/> | Otras:                           |                          |                                    |                          |

P2: Con que frecuencia viene usted a este establecimiento:

[Respuesta única]

|                       |                          |                           |                          |                |                          |
|-----------------------|--------------------------|---------------------------|--------------------------|----------------|--------------------------|
| Sólo de vez en cuando | <input type="checkbox"/> | Compra habitual / semanal | <input type="checkbox"/> | Compra mensual | <input type="checkbox"/> |
|-----------------------|--------------------------|---------------------------|--------------------------|----------------|--------------------------|

P3.- ¿Conoce, aunque no los haya comprado nunca, productos de alimentación y limpieza con Marca de Distribuidor, también conocidos como Marcas Blancas? (por ejemplo El Corte Inglés, Carrefour, Día, Hacendado, Eroski, Caprabo, Super Sol, etc).

[Respuesta única]

|    |                          |    |                          |
|----|--------------------------|----|--------------------------|
| SI | <input type="checkbox"/> | NO | <input type="checkbox"/> |
|----|--------------------------|----|--------------------------|



P4.- De los siguientes productos, podría indicarnos cuanto compra marcas de Fabricante (como Nestlé, Sos, Danone, Cuca, etc), y cuanto Marcas de Distribuidor (Marcas Blancas como Mercadona, El Corte Inglés, Carrefour, Dia, Eroski, Caprabo, Supersol, etc).

**[Mostrar el cuadro y explicar para que señale dónde marcaría para cada producto.]**

Si ha comprado Marcas de Distribuidor señálelo, si puede indicando la marca concreta

| [Marque con X<br>la casilla correspondiente] | Respuesta única por fila]            |                                  |                                    |                                    |  | ¿Ha Comprado hoy<br>Marca de Distribuidor? |    |       |
|--|--------------------------------------|----------------------------------|------------------------------------|------------------------------------|--|--|----|-------|
|  | Siempre<br>Marcas<br>Fabri-<br>cante | Más<br>Marcas<br>Fabri-<br>cante | Igual<br>unas<br>marcas y<br>otras | Más<br>Marcas<br>Distri-<br>buidor | Siempre<br>Marcas<br>Distri-<br>buidor | SI   | NO | MARCA |
|  | Postres y preparados lácteos         |                                  |                                    |                                    |  |  |    |       |
| Leche  |                                      |                                  |                                    |                                    |  |  |    |       |
| Aceites                                      |                                      |                                  |                                    |                                    |  |  |    |       |
| Cafés  |                                      |                                  |                                    |                                    |  |  |    |       |
| Chocolates                                   |                                      |                                  |                                    |                                    |  |  |    |       |
| Mermeladas y miel                            |                                      |                                  |                                    |                                    |  |  |    |       |
| Galletas y pan molde                         |                                      |                                  |                                    |                                    |  |  |    |       |
| Legumbres cocidas                            |                                      |                                  |                                    |                                    |  |  |    |       |
| Conservas pescado                            |                                      |                                  |                                    |                                    |  |  |    |       |
| Conservas vegetales                          |                                      |                                  |                                    |                                    |  |  |    |       |
| Pastas                                       |                                      |                                  |                                    |                                    |  |  |    |       |
| Bebidas (refrescos)                          |                                      |                                  |                                    |                                    |  |  |    |       |
| Cervezas                                     |                                      |                                  |                                    |                                    |  |  |    |       |
| Vinos / licores                              |                                      |                                  |                                    |                                    |  |  |    |       |
| Turrones                                     |                                      |                                  |                                    |                                    |  |  |    |       |
| Platos preparados, pizzas.                   |                                      |                                  |                                    |                                    |  |  |    |       |
| Detergentes                                  |                                      |                                  |                                    |                                    |  |  |    |       |
| Limpiahogares                                |                                      |                                  |                                    |                                    |  |  |    |       |
| Gel / Champú / pasta de dientes              |                                      |                                  |                                    |                                    |  |  |    |       |

P5.- A continuación le voy a leer una serie de razones que pueden influirle a la hora de decidirse a comprar Marcas de Distribuidor. Por favor, díganos lo importantes que son para usted.

| [Respuesta única por fila]                      | Nada importante <span style="float: right;">Muy importante</span> |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|
|   | 1   | 2 | 3 | 4 | 5 | 6 | 7 |
| Ofertas   |   |   |   |   |   |   |   |
| Publicidad en Tv, radio, prensa.                |   |   |   |   |   |   |   |
| Publicidad, carteles degustaciones en la tienda |   |   |   |   |   |   |   |
| Opiniones de amigos/ familiares.                |   |   |   |   |   |   |   |
| Confianza y garantía de Marcas del Distribuidor |   |   |   |   |   |   |   |
| Saber quien fabrica ese producto.               |   |   |   |   |   |   |   |
| Precio  |   |   |   |   |   |   |   |



P6.- Por favor, indique su grado de acuerdo /desacuerdo con las siguientes afirmaciones sobre las Marcas de Distribuidor, siendo 1 nada de acuerdo y 7 totalmente de acuerdo.

|  | Nada de acuerdo |   |   |   | Total de acuerdo |   |   |
|--|-----------------|---|---|---|------------------|---|---|
| [Respuesta única por fila]   | 1               | 2 | 3 | 4 | 5                | 6 | 7 |
| Creo que las Marcas de Distribuidor son <b>tan buenas</b> como las Marcas de Fabricante                            |                 |   |   |   |                  |   |   |
| Creo que las Marcas de Distribuidor y las Marcas de Fabricante son <b>elaboradas por los mismos fabricantes</b>    |                 |   |   |   |                  |   |   |
| En relación con el año pasado, <b>compro más productos</b> con Marcas del Distribuidor                             |                 |   |   |   |                  |   |   |
| Me gusta comprar Marcas de Distribuidor.   |                 |   |   |   |                  |   |   |
| Las Marcas de distribuidor <b>me ayudan a moderar</b> mi gasto en la compra.                                       |                 |   |   |   |                  |   |   |
| Me gustaría mayor información <b>sobre el fabricante</b> de las MDD que compro.                                    |                 |   |   |   |                  |   |   |
| Tengo <b>confianza en las MDD</b> que compro porque <b>confío en la calidad del establecimiento</b> que las vende. |                 |   |   |   |                  |   |   |
| El <b>precio</b> de las Marcas de Distribuidor es una de las principales razones para comprarlas                   |                 |   |   |   |                  |   |   |
| El <b>envase</b> de las Marcas Distribuidor es una de las principales razones para comprarlas                      |                 |   |   |   |                  |   |   |
| La <b>calidad</b> de las Marcas Distribuidor es una de las principales razones para comprarlas                     |                 |   |   |   |                  |   |   |
| La <b>confianza</b> (garantía) en las Marcas de Distribuidor es una de las principales razones para comprarlas     |                 |   |   |   |                  |   |   |

P7.- Señale la nota que pondría a las Marcas de Distribuidor de este establecimiento / tienda, para cada una de las siguientes cualidades que le voy a leer, siendo 0 la peor nota y 10 la mejor:

|                                     | Muy mala |   |   |   |   | Muy buena |   |   |   |   |    |
|-------------------------------------|----------|---|---|---|---|-----------|---|---|---|---|----|
| [Respuesta única por fila]          | 0        | 1 | 2 | 3 | 4 | 5         | 6 | 7 | 8 | 9 | 10 |
| Calidad                             |          |   |   |   |   |           |   |   |   |   |    |
| Precio                              |          |   |   |   |   |           |   |   |   |   |    |
| Envase                              |          |   |   |   |   |           |   |   |   |   |    |
| Confianza / Garantía                |          |   |   |   |   |           |   |   |   |   |    |
| Sabor (en alimentos)                |          |   |   |   |   |           |   |   |   |   |    |
| Eficacia (detergentes, limpiadores) |          |   |   |   |   |           |   |   |   |   |    |



P8.- Datos de clasificación:

[Respuesta única]

|                            |           |       |              |        |           |  |   |  |   |  |
|----------------------------|-----------|-------|--------------|--------|-----------|--|---|--|---|--|
| a) Sexo:                   |           | MUJER |              | HOMBRE |           |  |   |  |   |  |
| b) Edad:                   | < 25 años |       | 25 / 50 años |        | > 50 años |  |   |  |   |  |
| c) Residencia              |           | CERCA |              | LEJOS  |           |  |   |  |   |  |
| d) Nº personas en el hogar |           |       |              |        |           |  |   |  |   |  |
| e) Status socioeconómico   | A         |       | B            |        | C         |  | D |  | E |  |

| 1. PROFESION (*)                            |   | 2. ESTUDIOS (*)         |  |                   |                    |              |
|---|---|-------------------------|--|-------------------|--------------------|--------------|
|   |   | ESTUDIOS UNIVERSITARIOS | ESTUDIOS POSTERIORES A ENSEÑANZAS MEDIAS | ENSEÑANZAS MEDIAS | ESTUDIOS PRIMARIOS | SIN ESTUDIOS |
| TRABAJADOR POR CUENTA PROPIA                | EMPRESARIO CON 6 O MÁS EMPLEADOS                      | A                       | A  | A                 | B                  | C            |
|   | EMPRESARIO CON 5 O MENOS EMPLEADOS                    | B                       | C  | C                 | D                  | D            |
|   | AGRICULTOR (GRAN EXPLOTACION)                         | B                       | C  | C                 | C                  | D            |
|   | AGRICULTOR (PEQUEÑA EXPLOTACION)                      | C                       | C  | D                 | D                  | E            |
| POR CUENTA PROPIA O AJENA                   | PROFESIONALES   | B                       | C  | C                 | C                  | D            |
| POR CUENTA AJENA (SECTOR PUBLICO O PRIVADO) | DIRECTIVOS CON 6 O MÁS SUBORDINADOS                   | A                       | A  | B                 | C                  | C            |
|   | DIRECTIVOS CON 5 O MENOS SUBORDINADOS                 | A                       | B  | B                 | C                  | C            |
|   | MANDOS INTERMEDIOS                                    | B                       | B  | C                 | C                  | D            |
|   | OTROS EMPLEOS EN OFICINAS                             | B                       | C  | C                 | C                  | D            |
|   | OTROS EMPLEOS FUERA DE OFICINA Y OBREROS CUALIFICADOS | C                       | C  | C                 | D                  | D            |
|   | TRABAJADORES MANUALES NO CUALIFICADOS                 | C                       | D  | D                 | E                  | E            |
| NO TRABAJA                                  | JUBILADOS/ PENSIONISTAS                               | C                       | D  | D                 | D                  | E            |
|   | PARADOS   | C                       | C  | D                 | D                  | E            |

(\*) DE LA PERSONA CABEZA DE FAMILIA O PERSONA QUE APORTE MÁS INGRESOS A LA UNIDAD FAMILIAR

|               |   |               |
|---------------|---|---------------|
| CLASE SOCIAL: | A | ACOMODADA     |
|               | B | MEDIA - ALTA  |
|               | C | MEDIA - MEDIA |
|               | D | MEDIA - BAJA  |
|               | E | MODESTA       |